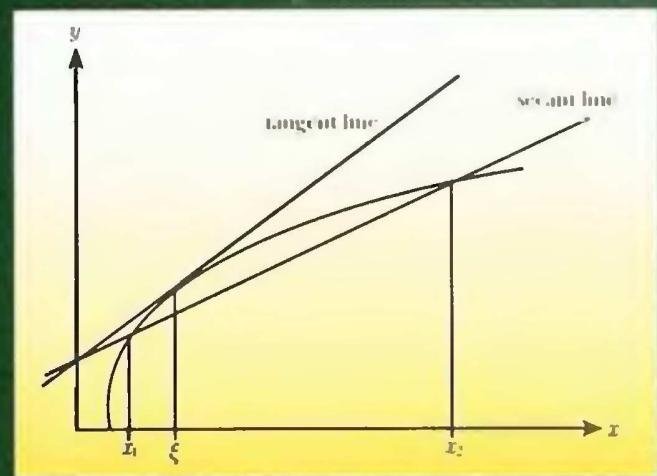
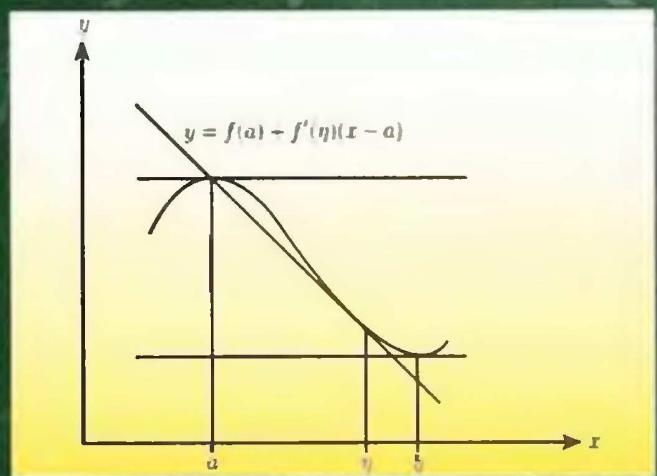
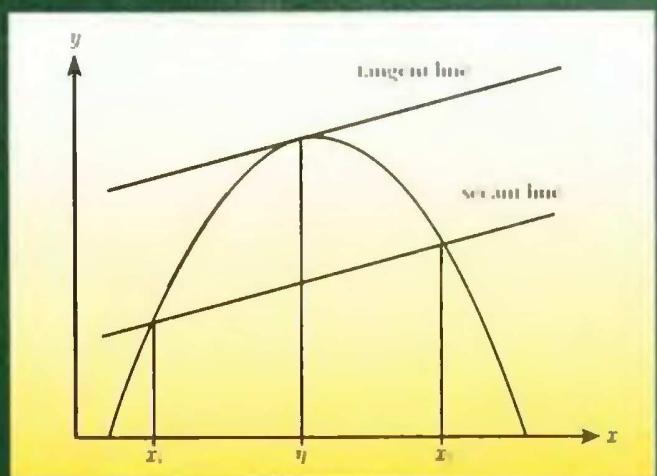


MEAN VALUE THEOREMS AND FUNCTIONAL EQUATIONS



P. K. Sahoo
T. Riedel

World Scientific

**MEAN VALUE
THEOREMS
AND
FUNCTIONAL
EQUATIONS**

**MEAN VALUE
THEOREMS
AND
FUNCTIONAL
EQUATIONS**

**P. K. Sahoo
T. Riedel**

*University of Louisville
Louisville, KY, USA*

 **World Scientific**
Singapore • New Jersey • London • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.
P O Box 128, Farrer Road, Singapore 912805
USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

MEAN VALUE THEOREMS AND FUNCTIONAL EQUATIONS

Copyright © 1998 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN 981-02-3544-5

This book is printed on acid-free paper.

Printed in Singapore by Uto-Print

To
Janos Aczél,
Palaniappan Kannappan and Berthold Schweizer
from whom we have learned so much about functional equations
and
to our families

Prasanna Sahoo
Thomas Riedel

Preface

Lagrange's mean value theorem is a very important result in analysis. It originated from Rolle's theorem, which was proved by the French mathematician Michel Rolle (1652-1719) for polynomials in 1691. This theorem appeared for the first time in the book *Methode pour resoudre les égalitez* without a proof and without any special emphasis. Rolle's theorem got its recognition when Joseph Lagrange (1736-1813) presented his mean value theorem in his book *Theorie des functions analytiques* in 1797. It received further recognition when Augustin Louis Cauchy (1789-1857) proved his mean value theorem in his book *Equationnes differentielles ordinaires*. Most of the results in Cauchy's book were established using the mean value theorem or indirectly Rolle's theorem. Since the discovery of Rolle's theorem (or Lagrange's mean value theorem), many papers have appeared directly or indirectly dealing with Rolle's theorem. Recently, many functional equations were studied motivated by various mean value theorems, and the main goal of this book is to present some results related to the mean value theorem (MVT) and its generalizations and many of these functional equations. This book started out as lecture notes for a seminar on mean value theorems and related functional equations held at the University of Louisville. The intent was to introduce advanced undergraduate and beginning graduate students to functional equations and to introduce those who had not yet taken a course in mathematical analysis to the basics of reading and writing proofs. We found that the mean value theorem with all it's generalizations and related functional equations was the ideal theme, a vast amount of material not usually included in a standard curriculum could be covered without spending large amounts of time on prerequisites.

Of course, we do not claim that this book covers everything related to the mean value theorem, there are many results we do not cover because they lie outside the scope of this book. Also, all functions appearing in the functional equations treated in this book are real or complex valued; we did not cover any functional equation where the unknown functions take on values on an algebraic structure such as group, ring or field. Our primary reason for writing this book is to introduce to our reader the simplicity and beauty of functional equations on one hand and the importance of the mean value theorem on the other hand. Almost all functional equations treated in this book are very elementary. The innovation of solving these functional equations lies in finding the right tricks for a particular case. One should not be misled by the elementary nature of the functional equations discussed in this book. We have stated many open problems related to some simple looking functional equations. It looks as though anyone could solve them, but, to the best of our knowledge, nobody has succeeded in finding the general solutions of these equations without any regularity conditions on the unknown functions.

We now give a brief description of the contents. Chapter 1 gives an account of additive and biadditive functions. In this chapter, we treat the Cauchy functional equation and show that continuous or locally integrable additive functions are linear. We further explore the behavior of discontinuous additive functions and show that they display a very strange behavior: their graphs are dense in the plane. To this end, we briefly discuss the Hamel basis and its use for constructing discontinuous additive functions. Additive functions on the real and complex planes are also treated in this chapter and, finally, we give a brief exposition on biadditive functions. This chapter concludes with a discussion of open problems.

In Chapter 2, Lagrange's mean value theorem is treated. Many examples are given to illustrate its importance. The mean value theorem is the motivation for many functional equations, some of which are treated in this chapter. Many of these results on functional equations are very recent. Further, we briefly describe the mean value theorem for divided differences and give some applications in defining the functional means. In this chapter we also examine the limiting behavior of the mean values. This chapter also concludes with a discussion of open problems.

Chapter 3 deals with a variation of Lagrange's mean value theorem due to Dimitri Pompeiu. Pompeiu's mean value theorem has been the source of motivations for many Stamate type functional equations. In this

chapter, we discuss several of them as well as a functional equation due to Marec Kuczma. We have simplified his proof by making some additional assumptions. This chapter examines also some functional equations which are extensions of previously studied equations and many are motivated by Simpson's rule for numerical integration. This chapter also concludes with a discussion of open problems.

In Chapter 4, we examine and extend Lagrange's mean value theorem to functions in two variables. We also discuss some functional equations that arise from such an extension. Some generalized mean value type functional equations and Cauchy's mean value theorem for functions in two variables are treated in this chapter as well. This chapter also concludes with a discussion of open problems.

Chapter 5 focuses on various generalizations of Lagrange's mean value theorem. We first treat all generalizations of the mean value theorem for functions from reals into reals. In this chapter, we examine various generalizations due to Flett, Trahan and many other mathematicians. We further treat mean value theorems for real valued functions on the plane, and present some results due to Clarke and Ledyaev. Mean value theorems for vector valued functions on reals are also treated in this chapter including results of McLeod and Sanderson. The mean value theorem for vector valued functions on the plane is also covered. A recent result of Furi and Martelli is included here and we also treat the mean value theorem for complex valued function on the complex plane. This chapter concludes with a conjecture due to Furi and Martelli and its recent proof by Ferrer.

Chapter 6 examines the mean value theorem and its generalizations for function with symmetric derivatives and Dini derivatives, respectively. Here, we introduce the notion of symmetric differentiation and then derive the mean value theorem for symmetrically differentiable functions. The notion of Dini derivatives is introduced with some well known examples. Finally, we present the mean value theorem for nondifferentiable functions.

Chapter 7 deals with the integral mean value theorem and its generalizations. Some applications of it are given, such as finding the integral representations of the arithmetic, geometric, logarithmic, and identric means and their extensions. Here we also discuss the iterations of arithmetic and geometric means and a theorem due to Kranz and Thews which states that if the mean values from the integral mean value theorem and differential mean value theorem occur at the same point then the underlying function is exponentially affine. This chapter concludes with some open problems.

This book would not have been completed without the support and encouragement from Robert Powers. He read a part of the manuscript and made many constructive criticisms for improvement. We are very grateful to him. We are also thankful to Elias Deeba for his many suggestions for improvement. We express our gratitude to Mario Martelli for providing us some valuable references and reprints of his papers related to Rolle's theorem. We are extremely thankful to Jürg Rätz for bringing our attention to his work on Cauchy's mean value theorem for divided differences. In this book, we have used results from many researchers and we have made honest efforts to pay credit to appropriate researchers. If we have missed anyone, we are sorry. If there is any resemblance with any published proof to a proof in this book it is because we could not improve the original proof of the author. We thank students who regularly attended our seminar on mean value theorems and functional equations, where the idea for writing this book was conceived. While this book was in its final form of revision, the first author visited a number of universities including the Technical University of Braunschweig (Germany), the University of Waterloo (Canada), the Aizu University (Japan), the Indian Statistical Institute (India) and the Sambalpur University (India) during his sabbatical in the Fall term of 1995. He would like to thank these universities for providing him with excellent work environment and hospitalities.

This book was typeset by authors in L^AT_EX, a macro package written by Leslie Lamport for Donald Knuth's T_EX typesetting package. The bibliography and index were compiled using BibTeX and MakeIndex, respectively.

Prasanna Sahoo
Thomas Riedel

Contents

Preface	vii
Chapter 1 Additive and Biadditive Functions	1
1.1 Continuous Additive Functions	1
1.2 Discontinuous Additive Functions	6
1.3 Other Criteria for Linearity	11
1.4 Additive Functions on the Real and Complex Plane	12
1.5 Biadditive Functions	18
1.6 Some Open Problems	22
Chapter 2 Lagrange's Mean Value Theorem and Related Functional Equations	25
2.1 Lagrange's Mean Value Theorem	25
2.2 Applications of the MVT	30
2.3 Associated Functional Equations	40
2.4 The MVT for Divided Differences	58
2.5 Limiting Behavior of Mean Values	64
2.6 Cauchy's MVT and Functional Equations	77
2.7 Some Open Problems	80
Chapter 3 Pompeiu's Mean Value Theorem and Associated Functional Equations	83
3.1 Pompeiu's Mean Value Theorem	83
3.2 Stamate Type Equations	85
3.3 An Equation of Kuczma	92

3.4 Equations Motivated by Simpson's Rule	98
3.5 Some Generalizations	108
3.6 Some Open Problems	125
Chapter 4 Two-dimensional Mean Value Theorems and Functional Equations	127
4.1 MVTs for Functions in Two Variables	127
4.2 Mean Value Type Functional Equations	129
4.3 Generalized Mean Value Type Equations	135
4.4 Cauchy's MVT for Functions in Two Variables	144
4.5 Some Open Problems	144
Chapter 5 Some Generalizations of Lagrange's Mean Value Theorem	147
5.1 MVTs for Real Functions	147
5.2 MVTs for Real Valued Functions on the Plane	156
5.3 MVTs for Vector Valued Functions on the Reals	160
5.4 MVTs for Vector Valued Functions on the Plane	163
5.5 MVTs for Functions on the Complex Plane	168
5.6 A Conjecture of Furi and Martelli	180
Chapter 6 Mean Value Theorems for Some Generalized Derivatives	181
6.1 Symmetric Differentiation of Real Functions	181
6.2 A Quasi-Mean Value Theorem	188
6.3 An Application	192
6.4 Generalizations of MVTs	193
6.5 Dini Derivatives of Real Functions	195
6.6 MVTs for Nondifferentiable Functions	201
Chapter 7 Some Integral Mean Value Theorems and Related Topics	207
7.1 The Integral MVT and Generalizations	207
7.2 Integral Representation of Means	218
7.3 Coincidence of Mean Values	224
7.4 Some Open Problems	227
Bibliography	233
Index	242

Chapter 1

Additive and Biadditive Functions

The goal of this chapter is to present some results concerning additive and biadditive functions. The study of additive functions dates back to A.M. Legendre who first attempted to determine the solution of the Cauchy functional equation

$$f(x + y) = f(x) + f(y)$$

for all $x, y \in \mathbb{R}$. The book of Kuczma (1985) gives an excellent exposition on additive functions. Additive functions have also found places in the books of Aczél (1966), Aczél (1987), Aczél and Dhombres (1989), and Smital (1988). The general solutions of many functional equations of two or more variables can be expressed in terms of additive, multiplicative, logarithmic or exponential functions. The functional equations, we shall be treating here, are only concerned with additive and biadditive functions and their Pexiderizations. In passing, we shall also examine solutions of some other equations related to the additive Cauchy equation. Some of the materials of this chapter are adapted from Aczél (1965) and Wilansky (1967).

1.1 Continuous Additive Functions

In this section, we define additive functions and then examine their behavior under various regularity assumptions such as continuity, differentiability, measurability, monotonicity.

Definition 1.1 A function $f : \mathbb{R} \rightarrow \mathbb{R}$, where \mathbb{R} is the set of real numbers, is called an *additive function* if and only if it satisfies the Cauchy functional

equation

$$f(x+y) = f(x) + f(y) \quad (1.1)$$

for all $x, y \in \mathbb{R}$.

The functional equation (1.1) was first treated by A.M. Legendre (1791) and C.F. Gauss (1809) but A.L. Cauchy (1821) first found its general continuous solution. The equation (1.1) has a privileged position in mathematics. It is encountered in almost all mathematical disciplines.

Definition 1.2 A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is called a *linear function* if and only if it is of the form

$$f(x) = mx \quad (\forall x \in \mathbb{R}),$$

where m is an arbitrary constant.

The graph of a linear function $f(x) = mx$ is a non-vertical line that passes through the origin and hence it is called linear. The only examples of additive functions which come readily to mind are linear functions. The question arises, are there any other additive function?

We begin by showing that the only continuous additive functions are those which are linear. This was the result proved by Cauchy in 1821.

Theorem 1.1 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous additive function. Then f is linear, that is, $f(x) = mx$ where m is an arbitrary constant.*

Proof: First, let us fix x and then we integrate both sides of (1.1) with respect to the variable y to get

$$\begin{aligned} f(x) &= \int_0^1 f(x) dy \\ &= \int_0^1 [f(x+y) - f(y)] dy \\ &= \int_x^{1+x} f(u) du - \int_0^1 f(y) dy, \quad \text{where } u = x + y. \end{aligned}$$

Since f is continuous, by using the Fundamental Theorem of calculus, we get

$$f'(x) = f(1+x) - f(x). \quad (1.2)$$

The additivity of f yields

$$f(1+x) = f(1) + f(x). \quad (1.3)$$

Substituting (1.3) into (1.2), we obtain

$$f'(x) = m,$$

where $m = f(1)$. Solving the above first order differential equation we obtain

$$f(x) = mx + c, \quad (1.4)$$

where c is an arbitrary constant. Letting the form of $f(x)$ from (1.4) in the functional equation (1.1), we see that $c = 2c$ and thus c must be zero. Therefore, from (1.4) we see that f is linear as asserted by the theorem. The proof of the theorem is now complete.

Notice that in Theorem 1.1, we use the continuity of f to conclude that f is also integrable. The integrability of f forced the additive function f to be linear. Thus every integrable additive function is also linear.

Definition 1.3 A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is said to be *locally integrable* if and only if it is integrable over every finite interval.

It is known that every locally integrable additive map is also linear. We give a short proof of this using an argument provided by Shapiro (1973). Assume f is locally integrable additive function. Hence $f(x+y) = f(x) + f(y)$ holds for all x and y in \mathbb{R} . From this and using the local integrability of f , we get

$$\begin{aligned} y f(x) &= \int_0^y f(x) dz \\ &= \int_0^y [f(x+z) - f(z)] dz \\ &= \int_x^{x+y} f(u) du - \int_0^y f(z) dz \\ &= \int_0^{x+y} f(u) du - \int_0^x f(u) du - \int_0^y f(u) du. \end{aligned}$$

The right side of the above equality is invariant under the interchange of x and y . Hence it follows that

$$y f(x) = x f(y)$$

for all $x, y \in \mathbb{R}$. Therefore, for $x \neq 0$, we obtain

$$\frac{f(x)}{x} = m,$$

where m is an arbitrary constant. This implies that $f(x) = mx$ for all $x \in \mathbb{R} \setminus \{0\}$. Since f is additive, we know that $f(0) = 0$. Together with this and the above, we conclude that f is a linear function in \mathbb{R} .

Although this proof of the above theorem is brief and involves only calculus, this proof is not very instructive. We will present now a different proof which will help us to understand the behavior of the additive functions a bit more. First, we prove a series of lemmas to reestablish the above theorem. We begin with the following definition.

Definition 1.4 A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is said to be *rationally homogeneous* if and only if

$$f(rx) = r f(x), \quad (1.5)$$

for all $x \in \mathbb{R}$ and all rational numbers r .

The following theorem shows that any additive function is rationally homogeneous.

Theorem 1.2 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be an additive function. Then f is rationally homogeneous. Moreover, f is linear on the set of rational numbers \mathbb{Q} .

Proof: Letting $x = 0 = y$ in (1.1) see that $f(0) = f(0) + f(0)$ and hence

$$f(0) = 0. \quad (1.6)$$

Substituting $y = -x$ in (1.1) and then using (1.6), we see that f is an odd function in \mathbb{R} , that is

$$f(-x) = -f(x) \quad (1.7)$$

for all $x \in \mathbb{R}$. Thus, so far, we have shown that an additive function is zero at the origin and it is an odd function. Next, we will show that an additive

function is rationally homogeneous. For any x ,

$$f(2x) = f(x+x) = f(x) + f(x) = 2f(x).$$

Hence

$$f(3x) = f(2x+x) = f(2x) + f(x) = 2f(x) + f(x) = 3f(x);$$

so in general (using induction)

$$f(nx) = n f(x) \quad (1.8)$$

for all positive integers n . If n is a negative integer, then $-n$ is a positive integer and by (1.8) and (1.7), we get

$$\begin{aligned} f(nx) &= f(-(-n)x) \\ &= -f(-nx) \\ &= -(-n)f(x) \\ &= n f(x). \end{aligned}$$

Thus, we have shown $f(nx) = n f(x)$ for all integers n and all $x \in \mathbb{R}$. Next, let r be an arbitrary rational number. Hence, we have

$$r = \frac{k}{\ell}$$

where k is an integer and ℓ is a natural number. Further, $kx = \ell(rx)$. Using the integer homogeneity of f , we obtain

$$k f(x) = f(kx) = f(\ell(rx)) = \ell f(rx)$$

that is

$$f(rx) = \frac{k}{\ell} f(x) = r f(x).$$

Thus, f is rationally homogeneous. Further, letting $x = 1$ in the above equation and defining $m = f(1)$, we see that

$$f(r) = m r$$

for all rational numbers $r \in \mathbb{Q}$. Hence, f is linear on the set of rational numbers and the proof is now complete.

Now we present the second proof of Theorem 1.1. Let f be additive and continuous on the set of reals. For any real number x there exists a

sequence $\{r_n\}$ of rational numbers with $r_n \rightarrow x$. Since, f is additive, by Theorem 1.2, f is linear on the set of rational numbers. That is

$$f(r_n) = m r_n$$

for all n . Now using the continuity of f , we get

$$\begin{aligned} f(x) &= f\left(\lim_{n \rightarrow \infty} r_n\right) \\ &= \lim_{n \rightarrow \infty} f(r_n) \\ &= \lim_{n \rightarrow \infty} m r_n \\ &= mx. \end{aligned}$$

Theorem 1.3 *If an additive function is continuous at a point, then it is continuous everywhere.*

Proof: Let f be continuous at t and let x be any arbitrary point. Hence, we have $\lim_{y \rightarrow t} f(y) = f(t)$. Next, we show that f is continuous at x . Consider

$$\begin{aligned} \lim_{y \rightarrow x} f(y) &= \lim_{y \rightarrow x} f(y - x + x - t + t) \\ &= \lim_{y \rightarrow x} f(y - x + t) + f(x - t) \\ &= \lim_{y-x+t \rightarrow t} f(y - x + t) + f(x - t) \\ &= f(t) + f(x - t) \\ &= f(t) + f(x) - f(t) \\ &= f(x). \end{aligned}$$

This proves that f is continuous at x and the arbitrariness of x implies f is continuous everywhere. The proof is complete.

1.2 Discontinuous Additive Functions

In the previous section, we have shown that continuous additive functions are linear. Even if we relax the continuity condition to continuity at a point, still additive functions are linear. For many years the existence of discontinuous additive functions was an open problem. Mathematicians could neither prove that every additive function is continuous, nor exhibit

an example of a discontinuous additive function. It was the German mathematician G. Hamel in 1905 who first succeeded in proving that there exist discontinuous additive functions.

Now we begin our exploration on non-linear additive functions. First, we show that non-linear additive functions display a very strange behavior.

Definition 1.5 The *graph of a function* $f : \mathbb{R} \rightarrow \mathbb{R}$ is the set

$$G = \{(x, y) \mid x \in \mathbb{R}, y = f(x)\}.$$

It is easy to note that the graph G of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ is subset of the plane \mathbb{R}^2 . The proof of our next theorem is similar to one found in Aczél (1987).

Theorem 1.4 *The graph of every non-linear additive function $f : \mathbb{R} \rightarrow \mathbb{R}$ is everywhere dense in the plane \mathbb{R}^2 .*

Proof: The graph G of f is given by

$$G = \{(x, y) \mid x \in \mathbb{R}, y = f(x)\}.$$

Choose a nonzero x_1 in \mathbb{R} . Since f is a non-linear additive map, for any constant m , there exists a nonzero real number x_2 such that

$$\frac{f(x_1)}{x_1} \neq \frac{f(x_2)}{x_2},$$

otherwise writing $m = \frac{f(x_1)}{x_1}$ and letting $x_1 = x$, we will have $f(x) = mx$ for all $x \neq 0$, and since $f(0) = 0$ this implies that f is linear contrary to our assumption that f is non-linear. This implies that

$$\begin{vmatrix} x_1 & f(x_1) \\ x_2 & f(x_2) \end{vmatrix} \neq 0,$$

so that the vectors $X_1 = (x_1, f(x_1))$ and $X_2 = (x_2, f(x_2))$ are linearly independent and thus they span the whole plane \mathbb{R}^2 . This means that for any vector $X = (x, f(x))$ there exist real numbers r_1 and r_2 such that

$$X = r_1 X_1 + r_2 X_2.$$

If we permit only rational numbers r_1, r_2 then, by their appropriate choice, we can get with $r_1 X_1 + r_2 X_2$ arbitrarily close to any given plane vector X

(since the rational numbers \mathbb{Q} are dense in reals \mathbb{R} and hence \mathbb{Q}^2 is dense in \mathbb{R}^2). Now,

$$\begin{aligned}\rho_1 X_1 + \rho_2 X_2 &= \rho_1(x_1, f(x_1)) + \rho_2(x_2, f(x_2)) \\ &= (\rho_1 x_1 + \rho_2 x_2, \rho_1 f(x_1) + \rho_2 f(x_2)) \\ &= (\rho_1 x_1 + \rho_2 x_2, f(\rho_1 x_1 + \rho_2 x_2)).\end{aligned}$$

Thus, the set

$$\widehat{G} = \{(x, y) \mid x = \rho_1 x_1 + \rho_2 x_2, y = f(\rho_1 x_1 + \rho_2 x_2), \rho_1, \rho_2 \in \mathbb{Q}\}$$

is everywhere dense in \mathbb{R}^2 . Since,

$$\widehat{G} \subset G,$$

the graph G of our non-linear additive function f is also dense in \mathbb{R}^2 . The proof of the theorem is now complete.

The graph of an additive continuous function is a straight line that passes through the origin. The graph of a non-linear additive function is dense in the plane. Next, we introduce the concept of Hamel basis to construct a discontinuous additive function.

Let us consider the set

$$S = \{s \in \mathbb{R} \mid s = u + v\sqrt{2} + w\sqrt{3}, u, v, w \in \mathbb{Q}\}$$

whose elements are rational linear combination of $1, \sqrt{2}, \sqrt{3}$. Further, this rational combination is unique. That is if an element $s \in S$ has two different rational linear combinations, for instance,

$$s = u + v\sqrt{2} + w\sqrt{3} = u' + v'\sqrt{2} + w'\sqrt{3}$$

then $u = u'$, $v = v'$ and $w = w'$. To prove this we note that this assumption implies that

$$(u - u') + (v - v')\sqrt{2} + (w - w')\sqrt{3} = 0.$$

Letting $a = (u - u')$, $b = (v - v')$ and $c = (w - w')$, we see that the above expression reduces to

$$a + b\sqrt{2} + c\sqrt{3} = 0.$$

Next, we show that $a = 0 = b = c$. The above expression yields

$$b\sqrt{2} + c\sqrt{3} = -a$$

and squaring both sides, we have

$$2bc\sqrt{6} = a^2 - 2b^2 - 3c^2.$$

This implies that b or c is zero; otherwise, we may divide both sides by $2bc$ and get

$$\sqrt{6} = \frac{a^2 - 2b^2 - 3c^2}{2bc}$$

contradicting the fact that $\sqrt{6}$ is an irrational number. If $b = 0$, then we have $a + c\sqrt{3} = 0$; this implies that $c = 0$ (else $\sqrt{3} = -\frac{a}{c}$, is a rational number contrary to the fact that $\sqrt{3}$ is an irrational number). Similarly if $c = 0$, we obtain that $b = 0$. Thus both b and c are zero. Hence it follows immediately that $a = 0$.

If we call

$$B = \{1, \sqrt{2}, \sqrt{3}\}$$

then every element of S is a *unique* rational linear combination of the elements of B . This set B is called a Hamel basis for the set S . Formally, a Hamel basis is defined as follows.

Definition 1.6 Let S be a set of real numbers and let B be a subset of S . Then B is called a *Hamel basis* for S if every member of S is a unique (finite) rational linear combination of B .

If the set S is the set of reals, then using the axiom of choice (or by transfinite induction) it can be shown that a Hamel basis B for \mathbb{R} exists. The proof of this is beyond the scope of this book.

There is a close connection between additive functions and Hamel bases. To exhibit an additive function it is sufficient to give its values on a Hamel basis, and these values can be assigned arbitrarily. This is the content of the next two theorems.

Theorem 1.5 Let B be a Hamel basis for \mathbb{R} . If two additive functions have the same value at each member of B , then they are equal.

Proof: Let f_1 and f_2 be two additive functions having the same value at each member of B . Then $f_1 - f_2$ is additive. Let us write $f = f_1 - f_2$. Let x be any real number. Then there are numbers b_1, b_2, \dots, b_n in B and rational numbers r_1, r_2, \dots, r_n such that

$$x = r_1 b_1 + r_2 b_2 + \dots + r_n b_n.$$

Hence

$$\begin{aligned}
 f_1(x) - f_2(x) &= f(x) \\
 &= f(r_1 b_1 + r_2 b_2 + \cdots + r_n b_n) \\
 &= f(r_1 b_1) + f(r_2 b_2) + \cdots + f(r_n b_n) \\
 &= r_1 f(b_1) + r_2 f(b_2) + \cdots + r_n f(b_n) \\
 &= r_1 [f_1(b_1) - f_2(b_1)] + r_2 [f_1(b_2) - f_2(b_2)] \\
 &\quad + \cdots + r_n [f_1(b_n) - f_2(b_n)] \\
 &= 0.
 \end{aligned}$$

Thus, we have $f_1 = f_2$ and the proof is complete.

Theorem 1.6 *Let B be a Hamel basis for \mathbb{R} . Let $g : B \rightarrow \mathbb{R}$ be an arbitrary function defined on B . Then there exists an additive function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $f(b) = g(b)$ for each $b \in B$.*

Proof: For each real number x there can be found b_1, b_2, \dots, b_n in B and rational numbers r_1, r_2, \dots, r_n with

$$x = r_1 b_1 + r_2 b_2 + \cdots + r_n b_n.$$

We define $f(x)$ to be

$$r_1 g(b_1) + r_2 g(b_2) + \cdots + r_n g(b_n).$$

This defines $f(x)$ for all x . This definition is unambiguous since, for each x , the choice of $b_1, b_2, \dots, b_n, r_1, r_2, \dots, r_n$ is unique, except for the order in which b_i and r_i are selected. For each b in B , we have $f(b) = g(b)$ by definition of f . Next, we show that f is additive on the reals. Let x and y be any two real numbers. Then

$$\begin{aligned}
 x &= r_1 a_1 + r_2 a_2 + \cdots + r_n a_n \\
 y &= s_1 b_1 + s_2 b_2 + \cdots + s_m b_m,
 \end{aligned}$$

where $r_1, r_2, \dots, r_n, s_1, s_2, \dots, s_m$ are rational numbers and $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_m$ are members of the Hamel basis B . The two sets $\{a_1, a_2, \dots, a_n\}$ and $\{b_1, b_2, \dots, b_m\}$ may have some members in common. Let the union of these two sets be $\{c_1, c_2, \dots, c_\ell\}$. Then $\ell \leq m + n$, and

$$\begin{aligned}
 x &= u_1 c_1 + u_2 c_2 + \cdots + u_\ell c_\ell \\
 y &= v_1 c_1 + v_2 c_2 + \cdots + v_\ell c_\ell,
 \end{aligned}$$

where $u_1, u_2, \dots, u_\ell, v_1, v_2, \dots, v_\ell$ are rational numbers, several of which may be zero. Now

$$x + y = (u_1 + v_1)c_1 + (u_2 + v_2)c_2 + \cdots + (u_\ell + v_\ell)c_\ell$$

and

$$\begin{aligned} f(x+y) &= f((u_1 + v_1)c_1 + (u_2 + v_2)c_2 + \cdots + (u_\ell + v_\ell)c_\ell) \\ &= (u_1 + v_1)g(c_1) + (u_2 + v_2)g(c_2) + \cdots + (u_\ell + v_\ell)g(c_\ell) \\ &= [(u_1g(c_1) + u_2g(c_2) + \cdots + u_\ell g(c_\ell))] \\ &\quad + [(v_1g(c_1) + v_2g(c_2) + \cdots + v_\ell g(c_\ell))] \\ &= f(x) + f(y). \end{aligned}$$

Hence f is additive on the set of real numbers \mathbb{R} and the proof of the theorem is now complete.

With the help of a Hamel basis, next we construct a non-linear additive function. Let B be the Hamel basis for the set of real numbers \mathbb{R} . Let $b \in B$ be any element of B . Define

$$g(x) = \begin{cases} 0 & \text{if } x \in B \setminus \{b\} \\ 1 & \text{if } x = b. \end{cases}$$

By Theorem 1.6, there exists an additive function $f : \mathbb{R} \rightarrow \mathbb{R}$ with $f(x) = g(x)$ for each $x \in B$. Note that this f can not be linear since for $x \in B$ and $x \neq b$, we have

$$0 = \frac{f(x)}{x} \neq \frac{f(b)}{b}.$$

Therefore f is a non-linear additive function.

We end this section with the following remark. No concrete example of a Hamel basis is known, we only know that it exists. The graph of a discontinuous additive function is not easy to draw as the set $\{f(x) | x \in \mathbb{R}\}$ is dense in \mathbb{R} .

1.3 Other Criteria for Linearity

We have seen that the graph of a non-linear additive function f is dense in the plane. That is, every circle contains a point, (x, y) such that $y = f(x)$. We have also seen that an additive function f becomes linear when one

imposes continuity on f . One can weaken this continuity condition to continuity at a point and still get f to be linear. In this section, we present some other mild regularity conditions that force an additive function to be linear.

Theorem 1.7 *If an additive function f is either bounded from one side or monotonic, then it is linear.*

Proof: Suppose f is not linear. Then by Theorem 1.4, the graph of f is dense in the plane. Since f is bounded from the above, for some constant M the additive function f satisfies

$$f(x) \leq M, \quad x \in \mathbb{R},$$

and the graph of f avoids the set $A = \{x \in \mathbb{R} \mid f(x) > M\}$. Therefore it can not be dense on the plane which is a contradiction. Hence contrary to our assumption, f is linear. The rest of the theorem can be established in a similar manner. Now the proof is complete.

Definition 1.7 A function f is said to be *multiplicative* if and only if $f(xy) = f(x)f(y)$ for all numbers x and y .

Theorem 1.8 *If an additive function f is also multiplicative, then it is linear.*

Proof: For any positive number x ,

$$f(x) = f(\sqrt{x} \cdot \sqrt{x}) = f(\sqrt{x}) \cdot f(\sqrt{x}) = [f(\sqrt{x})]^2 \geq 0.$$

Therefore f is bounded from below and hence by Theorem 1.7, we see that f is linear. This completes the proof.

1.4 Additive Functions on the Real and Complex Plane

In this section, first we present some results concerning additive functions on the plane \mathbb{R}^2 and then study additive complex-valued functions on the complex plane. We begin this section with the following result.

Theorem 1.9 *If $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is additive on the plane \mathbb{R}^2 , then there exist additive functions $A_1, A_2 : \mathbb{R} \rightarrow \mathbb{R}$ such that*

$$f(x_1, x_2) = A_1(x_1) + A_2(x_2) \tag{1.9}$$

for all $x_1, x_2 \in \mathbb{R}$.

Proof: Let $x = (x_1, x_2)$ and $y = (y_1, y_2)$ be any two arbitrary points in the plane. The additivity of f yields

$$f(x + y) = f(x) + f(y)$$

that is

$$f(x_1 + y_1, x_2 + y_2) = f(x_1, x_2) + f(y_1, y_2).$$

We define $A_1(x_1) = f(x_1, 0)$ and $A_2(x_2) = f(0, x_2)$ and claim that A_1 and A_2 are additive. This can be seen from

$$\begin{aligned} A_1(x_1 + y_1) &= f(x_1 + y_1, 0) \\ &= f(x_1 + y_1, 0 + 0) \\ &= f(x_1, 0) + f(y_1, 0) \\ &= A_1(x_1) + A_1(y_1). \end{aligned}$$

Hence A_1 is additive on \mathbb{R} . Similarly, it can be shown that A_2 is also additive on \mathbb{R} . Next, we show that f is a superimposition of A_1 and A_2 . Note that $(x_1, x_2) = (x_1, 0) + (0, x_2)$ and

$$f(x_1, x_2) = f(x_1, 0) + f(0, x_2) = A_1(x_1) + A_2(x_2).$$

This completes the proof of the theorem.

The following theorem follows from Theorem 1.9 and Theorem 1.1.

Theorem 1.10 *If $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a continuous additive function on the plane \mathbb{R}^2 , then there exist constants c_1, c_2 such that*

$$f(x_1, x_2) = c_1 x_1 + c_2 x_2 \tag{1.10}$$

for all $x_1, x_2 \in \mathbb{R}$.

This result can be furthered strengthen by weakening the continuity of $f : \mathbb{R}^2 \rightarrow \mathbb{R}$.

Lemma 1.1 *If an additive function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is continuous with respect to each variable, then it is jointly continuous.*

Proof: Since the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is additive, by Theorem 1.9, we have

$$f(x, y) = A_1(x) + A_2(y)$$

for all $x, y \in \mathbb{R}$. Since f is continuous with respect to each variable, we see that A_1 and A_2 are continuous. Hence

$$\lim_{x \rightarrow x_o} A_1(x) = A_1(x_o) \quad \text{and} \quad \lim_{y \rightarrow y_o} A_2(y) = A_2(y_o).$$

In order to show f is jointly continuous, we compute

$$\begin{aligned} \lim_{(x,y) \rightarrow (x_o,y_o)} f(x,y) &= \lim_{(x,y) \rightarrow (x_o,y_o)} [A_1(x) + A_2(y)] \\ &= \lim_{x \rightarrow x_o} A_1(x) + \lim_{y \rightarrow y_o} A_2(y) \\ &= A_1(x_o) + A_2(y_o) = f(x_o, y_o). \end{aligned}$$

This shows that f is jointly continuous. Now this proof is complete.

In view of this result, one can force linearity in real-valued additive functions on the plane by assuming continuity in each variable. Further, one can extend Theorem 1.9 to real-valued additive functions on \mathbb{R}^n and the proof is left to the reader.

Theorem 1.11 *If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuous additive function on \mathbb{R}^n , then there exist constants c_1, c_2, \dots, c_n such that*

$$f(x_1, x_2, \dots, x_n) = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n \quad (1.11)$$

for all $x_1, x_2, \dots, x_n \in \mathbb{R}$.

In the remaining portion of this section, we examine complex-valued additive functions on the complex plane. We begin with a brief introduction to complex number system. Numbers of the form $a + b\sqrt{-1}$, where a and b are real numbers are called complex numbers. In the early 16th century, Cardan (1501-1576) worked with complex numbers in solving quadratic and cubic equations. In the 18th century, functions involving complex numbers were found by Euler. For a long time complex numbers enjoyed a poor reputation and were not generally considered legitimate numbers until the middle of the 19th century. Descartes, rejected complex roots of equations and named them “imaginary”. Euler also felt that complex numbers “exist only in the imagination” and considered complex roots of an equation useful only in showing that the equation actually has no solutions. Gauss gave a geometric representation to complex numbers and realized it was erroneous to assume “that there was some dark mystery in these numbers”. Today, complex numbers are widely accepted due to the work of Gauss. The formal definition of complex numbers was given by William Hamilton.

Definition 1.8 The *complex number system* \mathbb{C} is the set of ordered pairs of real numbers (x, y) with addition and multiplication defined by

$$\begin{aligned}(x, y) + (u, v) &= (x + u, y + v) \\ (x, y)(u, v) &= (xu - yv, xv + yu)\end{aligned}$$

for all $x, y, u, v \in \mathbb{R}$.

The thinking of a real number as either x or $(x, 0)$ and letting i denote the purely imaginary number $(0, 1)$, we can rewrite the following expression

$$(x, y) = (x, 0) + (0, 1)(y, 0)$$

as

$$(x, y) = x + iy.$$

If we denote the left side of the above representation by z , then we have $z = x + iy$. The real number x is called the *real part* of z and is denoted by $\operatorname{Re} z$. Similarly, the real number y is called the *imaginary part* of z and is denoted by $\operatorname{Im} z$. If z is a complex number of the form $x + iy$, then the complex number $x - iy$ is called the *conjugate* of z and is denoted by \bar{z} .

An arbitrary function $f : \mathbb{C} \rightarrow \mathbb{C}$ can be written as

$$f(z) = f_1(z) + i f_2(z), \quad (1.12)$$

where $f_1 : \mathbb{C} \rightarrow \mathbb{R}$ and $f_2 : \mathbb{C} \rightarrow \mathbb{R}$ are given by

$$f_1(z) = \operatorname{Re} f(z), \quad \text{and} \quad f_2(z) = \operatorname{Im} f(z). \quad (1.13)$$

If f is additive, then by (1.12) and (1.13) we have

$$\begin{aligned}f_1(z_1 + z_2) &= \operatorname{Re} f(z_1 + z_2) \\ &= \operatorname{Re} [f(z_1) + f(z_2)] \\ &= \operatorname{Re} f(z_1) + \operatorname{Re} f(z_2) \\ &= f_1(z_1) + f_1(z_2),\end{aligned}$$

and

$$\begin{aligned}f_2(z_1 + z_2) &= \operatorname{Im} f(z_1 + z_2) \\ &= \operatorname{Im} [f(z_1) + f(z_2)] \\ &= \operatorname{Im} f(z_1) + \operatorname{Im} f(z_2) \\ &= f_2(z_1) + f_2(z_2).\end{aligned}$$

Theorem 1.12 *If $f : \mathbb{C} \rightarrow \mathbb{C}$ is additive, then there exist additive functions $f_{kj} : \mathbb{R} \rightarrow \mathbb{R}$ ($k, j = 1, 2$) such that*

$$f(z) = f_{11}(Re z) + f_{12}(Im z) + i f_{21}(Re z) + i f_{22}(Im z).$$

Proof: By (1.12), we obtain

$$f(z) = f_1(z) + i f_2(z),$$

where $f_1 : \mathbb{C} \rightarrow \mathbb{R}$ and $f_2 : \mathbb{C} \rightarrow \mathbb{R}$ are real-valued functions on the complex plane. Since f is an additive function, f_1 and f_2 are also additive functions. Since the functions f_1 and f_2 can be considered as functions from \mathbb{R}^2 into \mathbb{R} , applying Theorem 1.9, we have the asserted result.

Our next theorem concerns the form of complex-valued continuous additive functions on the complex plane.

Theorem 1.13 *If $f : \mathbb{C} \rightarrow \mathbb{C}$ is a continuous additive function, then there exist complex constants c_1 and c_2 such that*

$$f(z) = c_1 z + c_2 \bar{z} \tag{1.14}$$

where \bar{z} denotes the complex conjugate of z .

Proof: Since f is additive, by Theorem 1.12, we get

$$f(z) = f_{11}(Re z) + f_{12}(Im z) + i f_{21}(Re z) + i f_{22}(Im z),$$

where $f_{kj} : \mathbb{R} \rightarrow \mathbb{R}$ ($k, j = 1, 2$) are real-valued additive functions on the reals. The continuity of f implies the continuity of each function f_{kj} and hence

$$f_{kj}(x) = c_{kj} x$$

where c_{kj} ($k, j = 1, 2$) are real constants. Thus, using the form of $f(z)$ and the form of f_{kj} , we get

$$\begin{aligned}
 f(z) &= c_{11} \operatorname{Re} z + c_{12} \operatorname{Im} z + i c_{21} \operatorname{Re} z + i c_{22} \operatorname{Im} z \\
 &= (c_{11} + i c_{21}) \operatorname{Re} z + (c_{12} + i c_{22}) \operatorname{Im} z \\
 &= a \operatorname{Re} z + b \operatorname{Im} z \quad \text{where } a = c_{11} + i c_{21}, b = c_{21} + i c_{22} \\
 &= a \operatorname{Re} z - i(bi) \operatorname{Im} z \\
 &= \frac{a+bi}{2} \operatorname{Re} z + \frac{a-bi}{2} \operatorname{Re} z - \frac{a+bi}{2} i \operatorname{Im} z + \frac{a-bi}{2} i \operatorname{Im} z \\
 &= \frac{a-bi}{2} \operatorname{Re} z + \frac{a-bi}{2} i \operatorname{Im} z + \frac{a+bi}{2} \operatorname{Re} z - \frac{a+bi}{2} i \operatorname{Im} z \\
 &= \frac{a-bi}{2} (\operatorname{Re} z + i \operatorname{Im} z) + \frac{a+bi}{2} (\operatorname{Re} z - i \operatorname{Im} z) \\
 &= \frac{a-bi}{2} z + \frac{a+bi}{2} \bar{z} \\
 &= c_1 z + c_2 \bar{z},
 \end{aligned}$$

where $c_1 = \frac{a-bi}{2}$ and $c_2 = \frac{a+bi}{2}$ are complex constants. This completes the proof of the theorem.

Note that unlike the real-valued continuous additive functions on the reals, the complex-valued continuous additive functions on the complex plane are not linear. The linearity can be restored if one assumes stronger regularity condition like analyticity instead of continuity.

Definition 1.9 A function $f : \mathbb{C} \rightarrow \mathbb{C}$ is said to be *analytic* if and only if f is differentiable on \mathbb{C} .

Theorem 1.14 If $f : \mathbb{C} \rightarrow \mathbb{C}$ is an analytic additive function, then there exists complex constant c such that

$$f(z) = cz,$$

that is f is linear.

Proof: Since f is analytic, it is differentiable. Differentiating

$$f(z_1 + z_2) = f(z_1) + f(z_2) \tag{1.15}$$

with respect to z_1 , we get

$$f'(z_1 + z_2) = f'(z_1)$$

for all z_1 and z_2 in \mathbb{C} . Hence, letting $z_1 = 0$ and $z_2 = z$, we get

$$f'(z) = c,$$

where $c = f'(0)$ is a complex constant. From the above, we see that

$$f(z) = cz + b,$$

where b is a complex constant. Inserting this form of $f(z)$ into (1.15), we obtain $b = 0$ and hence the asserted solution follows. This completes the proof of the theorem.

We close this section with the following remark. It is surprising that Theorem 1.8 fails for complex-valued functions on the complex plane. It is well known that there is a discontinuous automorphism of the complex plane (see Kamke (1927)). An *automorphism* is a map that is one-to-one, onto, additive and multiplicative.

1.5 Biadditive Functions

In this section of this chapter, we examine biadditive functions. We begin this section with the following definition of the biadditive function.

Definition 1.10 A function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is said to be *biadditive* if and only if it is additive in each variable, that is

$$f(x + y, z) = f(x, z) + f(y, z) \tag{1.16}$$

$$f(x, y + z) = f(x, y) + f(x, z) \tag{1.17}$$

for all $x, y, z \in \mathbb{R}$.

The only examples of biadditive functions which come readily to mind are a multiple of the product of independent variables. Thus if m is a constant and we define f by

$$f(x, y) = mxy, \quad x, y \in \mathbb{R}$$

then f is biadditive. The question arises, are there any other biadditive functions?

Theorem 1.15 *Every continuous biadditive map $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is of the form*

$$f(x, y) = mxy$$

for all $x, y \in \mathbb{R}$ and some arbitrary constant m in \mathbb{R} .

Proof: Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a continuous biadditive map. Hence f satisfies

$$f(x + y, z) = f(x, z) + f(y, z) \quad (1.18)$$

for all $x, y, z \in \mathbb{R}$. Letting $x = 0 = y$, in the above equation, we have the following condition

$$f(0, z) = 0 \quad (1.19)$$

for all $z \in \mathbb{R}$. For fixed z , defining $\phi(x) = f(x, z)$, we see that the equation (1.18) reduces to

$$\phi(x + y) = \phi(x) + \phi(y). \quad (1.20)$$

Since, f is continuous so also ϕ and therefore by (1.20), ϕ is linear, that is $\phi(x) = kx$; and

$$f(x, y) = k(y)x \quad (1.21)$$

where $k : \mathbb{R} \rightarrow \mathbb{R}$ is an arbitrary function. Since, f is additive in the second variable also, we have

$$f(x, y + z) = f(x, y) + f(x, z) \quad (1.22)$$

Letting (1.21) into (1.22), we have

$$xk(y + z) = xk(y) + xk(z)$$

for all $x, y, z \in \mathbb{R}$. If $x \neq 0$, then the above equation yields

$$k(y + z) = k(y) + k(z). \quad (1.23)$$

Again, using the additivity of f , we see that k is also additive and hence it is linear. Thus $k(y) = my$ for some arbitrary constant m . This in (1.21) yields

$$f(x, y) = mxy \quad (1.24)$$

for all $y \in \mathbb{R}$ and for all nonzero x in \mathbb{R} . If $x = 0$, then from (1.17), we see that $f(0, y) = 0$ and therefore (1.24) holds for all $x, y \in \mathbb{R}$. Now the proof is complete.

In the next theorem, we present a general representation for the biadditive function in terms of a Hamel basis.

Theorem 1.16 *Every biadditive map $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ can be represented as*

$$f(x, y) = \sum_{k=1}^n \sum_{j=1}^m \alpha_{kj} r_k s_j, \quad (1.25)$$

where

$$x = \sum_{k=1}^n r_k b_k, \quad y = \sum_{j=1}^m s_j b_j,$$

the r_k, s_j being rational while the b_j are elements of a Hamel basis B and the α_{kj} arbitrary depending upon b_k and b_j .

Proof: Let B be a Hamel basis for the set of reals \mathbb{R} . Then every real number x can be represented as

$$x = \sum_{k=1}^n r_k b_k \quad (1.26)$$

with $b_k \in \mathbb{R}$ and with rational coefficient r_k . Similarly, any other real number y can also be represented as

$$y = \sum_{j=1}^m s_j b_j \quad (1.27)$$

with $b_j \in \mathbb{R}$ and with rational coefficient s_j . Since f is biadditive

$$f(x_1 + x_2, y) = f(x_1, y) + f(x_2, y) \quad (1.28)$$

$$f(x, y_1 + y_2) = f(x, y_1) + f(x, y_2) \quad (1.29)$$

for all $x_1, x_2, y_1, y_2 \in \mathbb{R}$. From (1.28) and (1.29), using induction, we have

$$f\left(\sum_{k=1}^n x_k, y\right) = \sum_{k=1}^n f(x_k, y) \quad (1.30)$$

$$f\left(x, \sum_{k=1}^n y_k\right) = \sum_{k=1}^n f(x, y_k). \quad (1.31)$$

Letting $x_1 = x_2 = \dots = x_n = x$ and $y_1 = y_2 = \dots = y_n = y$, in (1.30) and (1.31) respectively, we get

$$f(nx, y) = n f(x, y) = f(x, ny). \quad (1.32)$$

From (1.32) with $t = \frac{m}{n} x$ (that is $nt = mx$), we get

$$n f(t, y) = f(nt, y) = f(mx, y) = m f(x, y)$$

or

$$f(t, y) = \frac{m}{n} f(x, y).$$

That is

$$f\left(\frac{m}{n} x, y\right) = \frac{m}{n} f(x, y). \quad (1.33)$$

Since f is biadditive, we see that

$$f(x, 0) = 0 = f(0, y) \quad (1.34)$$

for all $x, y \in \mathbb{R}$. Next, substituting $x_2 = -x_1 = x$ in (1.28) and using (1.34) we obtain

$$f(-x, y) = -f(x, y). \quad (1.35)$$

From (1.35) and (1.33) we conclude that (1.32) is valid for all rational numbers. The same argument applies to the second variable and so we have for all rational numbers r and all real x and y

$$f(rx, y) = r f(x, y) = f(x, ry). \quad (1.36)$$

Hence by (1.28), (1.29), (1.30), (1.31) and (1.36), we obtain

$$\begin{aligned}
 f(x, y) &= f\left(\sum_{k=1}^n r_k b_k, \sum_{j=1}^m s_j b_j\right) \\
 &= \sum_{k=1}^n r_k f\left(b_k, \sum_{j=1}^m s_j b_j\right) \\
 &= \sum_{k=1}^n \sum_{j=1}^m r_k s_j f(b_k, b_j) \\
 &= \sum_{k=1}^n \sum_{j=1}^m r_k s_j \alpha_{kj},
 \end{aligned}$$

where $\alpha_{kj} = f(b_k, b_j)$. This completes the proof of the theorem.

1.6 Some Open Problems

In this concluding section, we seek solutions of two problems related to additive functions on restricted domains. These and some additional problems were posed in Sahoo (1995). The first problem is the following: Find all functions $f :]0, 1[\rightarrow \mathbb{R}$ satisfying the functional equation

$$f(xy) + f(x(1-y)) + f(y(1-x)) + f((1-x)(1-y)) = 0 \quad (1.37)$$

for all $x, y \in]0, 1[$. This problem was stated as an open problem in Ebanks, Sahoo and Sander (1990). It should be noted that if $f(x) = 4A(x) - A(1)$, where A is an additive function on reals, then it satisfies the functional equation (1.37). If f is assumed to be continuous (or measurable), then Daroczy and Jarai (1979) have shown that $f(x) = 4ax - a$, where a is an arbitrary constant. Recently, Maksa (1993) has posed the following problem at the Thirtieth International Symposium on Functional Equations: Find all functions $f : [0, 1] \rightarrow \mathbb{R}$ satisfying the functional equation

$$(1-x-y)f(xy) = xf(y(1-x)) + yf(x(1-y)) \quad (1.38)$$

for all $x, y \in [0, 1]$. One can easily show that if f is a solution of (1.38), then f is skew symmetric about $\frac{1}{2}$, that is $f(x) = -f(1-x)$, and $f(0) = 0$. Further, it is easy to note that Maksa's equation (1.38) implies equation

(1.37). To see this replace x by $1-x$ in (1.38) and add the resulting equation to (1.38) to obtain

$$y [f(xy) + f(x(1-y)) + f(y(1-x)) + f((1-x)(1-y))] = 0$$

for all $x, y \in]0, 1]$. Since $f(0) = 0$, the above equation yields (1.37) for all $x, y \in [0, 1]$. Thus, the general solution of (1.37) will provide the general solution of (1.38). Utilizing the solution given by Daroczy and Jarai (1979) of the equation (1.37), it is easy to show that if f is continuous (or measurable or almost open), then all solutions of (1.38) are of the form $f(x) = 0$.

The second problem is the following: Find all functions $f :]0, 1[\rightarrow \mathbb{R}$ satisfying the functional equation

$$f(xy) + f((1-x)(1-y)) = f(x(1-y)) + f(y(1-x)) \quad (1.39)$$

for all $x, y \in]0, 1[$. Daroczy and Jarai (1979) have also found the measurable solution of this functional equation. They have shown that any measurable solution of (1.39) is of the form $f(x) = ax^2 - ax + b \log x + c$, where a, b and c are arbitrary constants. Equation (1.39) appears as a problem posed by Lajko in 1974 when (1.39) holds for all x and y in \mathbb{R} . Eliezer (1974) has determined the differentiable solution of Lajko's problem. Eliezer proved that if f is differentiable and satisfies (1.39) for all x and y in \mathbb{R} , then $f(x) = ax^2 - ax + c$, where a and c are arbitrary constants.

Now at the first sight the above two functional equations (1.37) and (1.39) seem harmless – it looks as though anyone could solve them, but nobody has succeeded finding all the general solutions of these equations without any regularity assumptions of the unknown function f .

Chapter 2

Lagrange's Mean Value Theorem and Related Functional Equations

In this chapter, we present the mean value theorem of differential calculus and some of its applications. Further, we discuss many functional equations that can be motivated using the mean value theorem. All the functional equations treated in this chapter are used in characterizing polynomials. In this chapter, we also examine the mean value theorem for divided differences and give some applications toward the study of means. Finally, we prove Cauchy's mean value theorem and point out various functional equations that can be motivated using this general theorem.

2.1 Lagrange's Mean Value Theorem

One of the most important theorems in differential calculus is Lagrange's mean value theorem. This theorem was first discovered by Joseph Louis Lagrange (1736-1813) but the idea of applying the Rolle's theorem to a suitably contrived auxiliary function was given by Ossian Bonnet (1819-1892). However, the first statement of the theorem appears in a paper of the renowned physicists André-Marie Ampère (1775-1836). It is well known that many results of classical real analysis are a consequence of the mean value theorem. The proof of Rolle's theorem is based on the following two results, which are ordinarily proved in a calculus course. We will merely state them here and refer the interested reader to a calculus book, for example Janusz (1994).

Proposition 2.1 *If a differentiable function $f : \mathbb{R} \rightarrow \mathbb{R}$, attains an extreme value at a point c in an open interval $]a, b[$, then $f'(c) = 0$.*

Proposition 2.2 *A continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ attains its extreme values on any closed and bounded interval $[a, b]$.*

We begin with Rolle's theorem, which is as follows:

Theorem 2.1 *If f is continuous on $[x_1, x_2]$ and differentiable on (x_1, x_2) , and $f(x_1) = f(x_2)$, then there exists a point $\eta \in (x_1, x_2)$ such that $f'(\eta) = 0$.*

Proof. Since f is continuous and $[x_1, x_2]$ is a closed bounded interval, then by Proposition 2.2 f attains its maximum and minimum value on this interval. If both of these occur at the end points x_1, x_2 , then maximum and minimum value are equal and the function is constant, hence $f'(\eta) = 0$ for all η in (x_1, x_2) . If this is not the case then one of the extreme values occurs at a point $\eta \in (x_1, x_2)$ and by Proposition 2.1, we have $f'(\eta) = 0$. The proof is now complete.

Thus Rolle's theorem can be geometrically interpreted as follows: If there is a horizontal secant line to the graph of f , then there is a horizontal tangent to the graph that is supported at a point between the two points of intersection of the secant line.

The other interpretation of Rolle's theorem is that between any two real zeros of a differentiable real function f lies at least one critical point of f (zero of its first derivative f').

Rolle's theorem is generalized by rotating the graph of f , which yields Lagrange's mean value theorem.

Theorem 2.2 *For every real valued function f differentiable on an interval I and for all pairs $x_1 \neq x_2$ in I , there exists a point η depending on x_1 and x_2 such that*

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = f'(\eta(x_1, x_2)). \quad (2.1)$$

Proof: The proof follows the idea that the mean value theorem is just a 'rotated version' of Rolle's theorem. We consider the function

$$h(x) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) + f(x_1).$$

This is the equation of the line intersecting the graph of f at $(x_1, f(x_1))$ and $(x_2, f(x_2))$. If we now define

$$g(x) = f(x) - h(x),$$

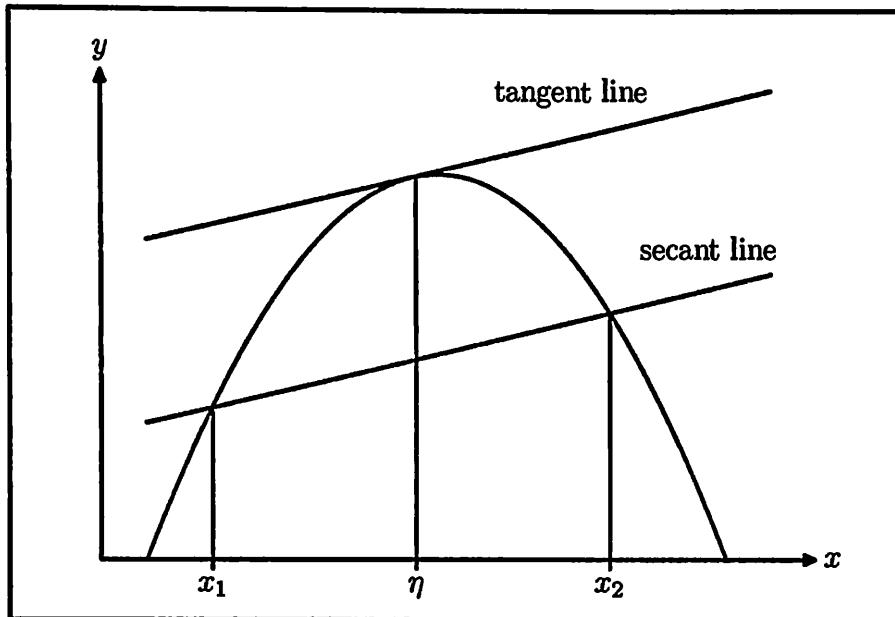


Fig. 2.1 A Geometrical Interpretation of the Mean Value Theorem.

then g is the result of rotating f and shifting it down to the x -axis. Since both f and h are continuous on $[x_1, x_2]$ and differentiable on $]x_1, x_2[$ so is g , and with a little algebra we find that $g(x_1) = g(x_2) = 0$ and thus g satisfies the hypotheses of Rolle's theorem. Now we may apply Rolle's theorem which results in the existence of an $\eta \in]x_1, x_2[$, where

$$0 = g'(\eta) = f'(\eta) - \frac{f(x_2) - f(x_1)}{x_2 - x_1},$$

and thus

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\eta).$$

The proof of the theorem is now complete.

The following is a pictorial proof of the Lagrange's mean value theorem. This proof was given by Swann (1997).

Intuitively, the assumptions of the theorem mean that the graph of the function $f(x)$ is smooth between $(x_1, f(x_1))$ and $(x_2, f(x_2))$ and has no sharp corners and vertical tangents since f is differentiable on the interval $]x_1, x_2[$. If the graph of $f(x)$ is not a straight line, some part of the graph will be above the line through $(x_1, f(x_1))$ and $(x_2, f(x_2))$ or below this line. Imagine a line that is parallel to the line through the points $(x_1, f(x_1))$ and

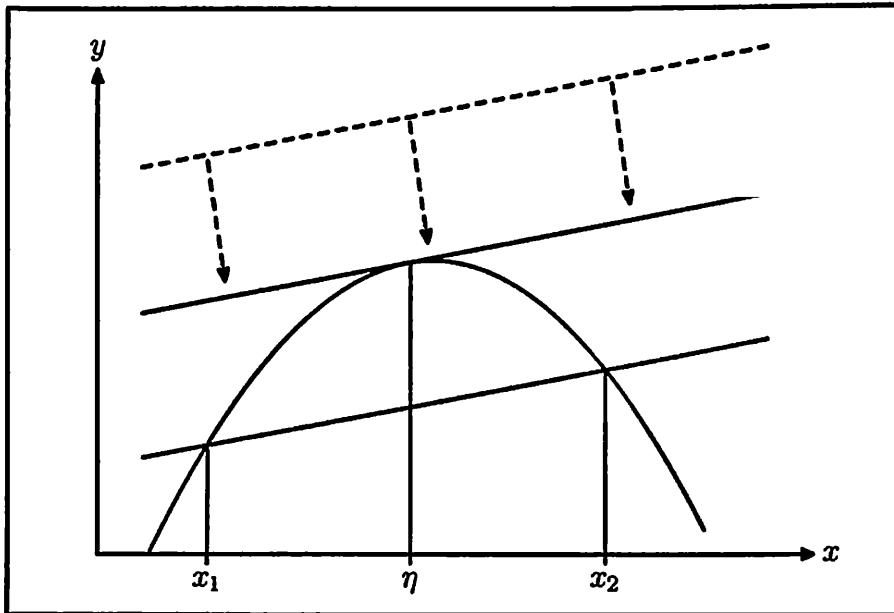


Fig. 2.2 A Graphical Proof of the Mean Value Theorem.

$(x_2, f(x_2))$ but far above the line. Move it down toward the line, keeping it parallel to the line through $(x_1, f(x_1))$ and $(x_2, f(x_2))$. Since there are no corners on the graph, when the line first hits the graph at some point $(\eta, f(\eta))$, surely it will be tangent to the graph at such a point. So, if our definition of the derivative as the slope of a line that is tangent to the graph at $(\eta, f(\eta))$ is any good, the slope of this tangent line must be $f'(\eta)$. But since this line is parallel to the line through $(x_1, f(x_1))$ and $(x_2, f(x_2))$, it will have the same slope as this line, that is

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\eta).$$

A similar argument holds if some of the graph is below the line. The above figure gives a geometrical illustration of this proof.

We close this section with another proof of Lagrange's theorem that does not use Proposition 2.1 and Proposition 2.2. This following proof is due to Tucker (1997) and Velleman (1998) (see Horn (1998)).

Start with a nonempty interval $[x_1, x_2]$ on which f is differentiable and define

$$m = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \quad \text{and} \quad y = \frac{x_2 - x_1}{2}.$$

Then y divides the interval $[x_1, x_2]$ into two subintervals of length $h = \frac{x_2 - x_1}{2}$. Observe that

$$\min\{m_1, m_2\} \leq m \leq \max\{m_1, m_2\},$$

where

$$m_1 = \frac{f(y) - f(x_1)}{h} \quad \text{and} \quad m_2 = \frac{f(x_2) - f(y)}{h}.$$

It follows by the intermediate value theorem, that the function

$$g(x) = \frac{f(x+h) - f(x)}{h}$$

takes the value m somewhere on $[a_1, b_1]$ such that

$$\frac{f(b_1) - f(a_1)}{b_1 - a_1} = m.$$

Iterate this reasoning to construct a nested sequence of intervals

$$[x_1, x_2] \supset [a_1, b_1] \supset [a_2, b_2] \supset \cdots \supset [a_n, b_n] \supset \cdots$$

such that

$$\frac{f(b_n) - f(a_n)}{b_n - a_n} = m$$

for all $n = 1, 2, \dots$ and $\lim_{n \rightarrow \infty} (b_n - a_n) = 0$. Let η be the unique point in the intersection of these intervals. If $\eta = a_N$ for some N , then $\eta = a_n$ for all $n > N$, so that

$$m = \frac{f(b_n) - f(\eta)}{b_n - \eta} \rightarrow f'(\eta).$$

Similarly, we get $m = f'(\eta)$ if $\eta = b_N$ for some N . If $a_n < \eta < b_n$ for all n , then

$$m = \mu_n \left[\frac{f(\eta) - f(a_n)}{\eta - a_n} \right] + (1 - \mu_n) \left[\frac{f(b_n) - f(\eta)}{b_n - \eta} \right]$$

for all n , where

$$0 < \mu_n = \frac{\eta - a_n}{b_n - a_n} < 1.$$

If both the quotients are within ϵ of $f'(\eta)$, then so is their convex combination, which means the m is within ϵ of $f'(\eta)$ for any $\epsilon > 0$. With a little

care, one can show that η is strictly between x_1 and x_2 . We leave this to the reader.

2.2 Applications of the MVT

The mean value theorem (MVT) has the following geometric interpretation. The tangent line to the graph of the function f at $\eta(x_1, x_2)$ is parallel to the secant line joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$. This is illustrated in Figure 2.1.

In this section we establish some results of differential and integral calculus using Lagrange's mean value theorem.

Lemma 2.1 *If $f'(x) = 0$ for all x in $]a, b[$, then f is a constant on $[a, b]$.*

Proof: Let x_1, x_2 be any two points in $]a, b[$, and suppose $f(x_1) \neq f(x_2)$, then, by the mean value theorem there is a $c \in]a, b[$ such that

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \neq 0,$$

contradicting the hypothesis that $f'(x) = 0$ for all $x \in]a, b[$.

An immediate consequence is the following:

Lemma 2.2 *If $f'(x) = g'(x)$ for all x in $]a, b[$, then f and g differ by a constant on $[a, b]$.*

Proof: Let $h(x) = f(x) - g(x)$, then $h'(x) = 0$, so by Lemma 2.1 $h(x) = c$, where c is a constant. Thus, f and g differ by a constant.

Lemma 2.3 *If $f'(x) > 0$ (< 0) for all x in $]a, b[$, then f is a strictly increasing (decreasing) function on $[a, b]$.*

Proof: Let $x_1 < x_2$ be in $[a, b]$, then by the mean value theorem there is a c in $]x_1, x_2[$, such that

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c) > 0$$

and since $x_2 - x_1 > 0$ we have $f(x_2) - f(x_1) > 0$ or $f(x_1) < f(x_2)$, and f is increasing.

An easy consequence of this is the following:

Lemma 2.4 *If $f''(x) > 0$ for all x in $]a, b[$, then f is a concave up on the interval $[a, b]$.*

The fundamental theorem of calculus, states that if f is a continuous function on $[a, b]$ and F is some indefinite integral of f on $[a, b]$, then

$$\int_a^b f(t) dt = F(b) - F(a). \quad (2.2)$$

This theorem can also be established by invoking the mean value theorem. Besides these theoretical applications, the mean value theorem has other applications. The next seven examples illustrate some other applications of the mean value theorem.

Example 2.1 The mean value theorem can be used to prove Bernoulli's inequality: If $x > -1$, then

$$(1 + x)^n \geq 1 + nx$$

for all $n \in \mathbb{R}$.

First, we suppose $x \geq 0$ and let $f(t) = (1 + t)^n$, for $t \in [0, x]$. Thus f satisfies the hypotheses of the mean value theorem and we have an $\eta \in]0, x[$ with

$$f(x) - f(0) = (x - 0) f'(\eta).$$

Thus we have

$$(1 + x)^n - 1 = x n (1 + \eta)^{n-1} \geq nx,$$

and hence $(1 + x)^n \geq 1 + nx$. The case when $-1 < x < 0$ can be handled similarly by considering $f(t) = (1 + t)^n$ for $t \in [x, 0]$.

Example 2.2 The mean value theorem can be used in proving the inequality

$$x \geq 1 + \ln x, \quad x > 0 \quad (2.3)$$

with equality if and only if $x = 1$.

To prove this, let $f(t) = \ln t$ for $t \in [1, b]$ and $b > 1$. The function f satisfies the hypothesis of the mean value theorem. Hence, there exists an $\eta \in]1, b[$ such that

$$f(b) - f(1) = (b - 1) f'(\eta)$$

which is

$$\ln b = \frac{b-1}{\eta}.$$

Since $\eta \in]1, b[$, we get

$$\frac{b-1}{b} < \frac{b-1}{\eta} < \frac{b-1}{1}$$

and thus

$$1 - \frac{1}{b} < \ln b < b - 1. \quad (2.4)$$

Using the left-hand inequality of (2.4), we obtain

$$\frac{1}{b} > 1 + \ln \left(\frac{1}{b} \right).$$

Since $b > 1$, it follows that $0 < \frac{1}{b} < 1$. Letting $x = \frac{1}{b}$, we have $x > 1 + \ln x$ if $0 < x < 1$. Next the right-hand side of the inequality (2.4) yields

$$\ln b < b - 1$$

for $b > 1$. This can be rewritten as

$$x > 1 + \ln x \quad \text{for } x > 1.$$

If $x = 1$, then clearly left side of (2.3) is equal to right side of (2.3). Thus, we have shown that for any $x > 0$, the inequality (2.3) holds with equality if and only if $x = 1$.

This inequality is widely used in information theory to establish the nonnegativity of the directed divergence. This inequality is further used for proving the arithmetic mean is greater than or equal to the geometric mean.

Example 2.3 The mean value theorem can be used in establishing the following inequality

$$a^\alpha < \{ a \alpha + b (1 - \alpha) \} b^{\alpha-1}, \quad (2.5)$$

for $0 < \alpha < 1$ and a, b are positive real numbers.

To prove the inequality (2.5), define a function f by

$$f(t) = t^\alpha, \quad t > 0, \quad 0 < \alpha < 1.$$

Then, evidently, f is continuous on $[a, b]$. Applying the mean value theorem to f , we obtain

$$\frac{f(b) - f(a)}{b - a} = f'(\eta)$$

for some η in the open interval $]a, b[$. This yields

$$\frac{b^\alpha - a^\alpha}{b - a} = \alpha \eta^{\alpha-1}. \quad (2.6)$$

Since $\eta \in]a, b[$, we obtain

$$\eta^{\alpha-1} > b^{\alpha-1}.$$

Hence, since $\alpha > 0$, we have

$$\alpha \eta^{\alpha-1} > \alpha b^{\alpha-1}.$$

Using (2.6) in the above inequality, we see that

$$b^\alpha - a^\alpha > (b - a) \alpha b^{\alpha-1}$$

which after some simplifications yields the inequality (2.5), that is

$$a^\alpha < \{a\alpha + b(1-\alpha)\} b^{\alpha-1}.$$

This inequality is used while proving the Hölder inequality in analysis.

Example 2.4 The mean value theorem can be used for showing that $(1 + \frac{1}{x})^x$ is an increasing function of x while $(1 + \frac{1}{x})^{x+1}$ is a decreasing function of x for $x > 0$.

Let us define a function f by

$$f(t) = \ln t, \quad t > 0.$$

Applying the mean value theorem to f , we get

$$f(x+1) - f(x) = f'(\eta)$$

for some $\eta \in]x, x+1[$. This yields

$$\ln(x+1) - \ln(x) = \frac{1}{\eta}, \quad x > 0. \quad (2.7)$$

Since

$$\begin{aligned}
 \frac{d}{dx} \left[\ln \left(1 + \frac{1}{x} \right)^x \right] &= \frac{d}{dx} [x (\ln(x+1) - \ln(x))] \\
 &= \ln(x+1) - \ln(x) + x \left[\frac{1}{x+1} - \frac{1}{x} \right] \\
 &= \ln(x+1) - \ln(x) - \frac{1}{x+1} \\
 &= \frac{1}{\eta} - \frac{1}{x+1} > 0, \quad \text{by (2.7),}
 \end{aligned}$$

and $\ln(x)$ is an increasing function, we conclude that $(1 + \frac{1}{x})^x$ is an increasing function of x .

To show that $(1 + \frac{1}{x})^{x+1}$ is a decreasing function of x we proceed in a similar manner and show

$$\begin{aligned}
 \frac{d}{dx} \left[\ln \left(1 + \frac{1}{x} \right)^{x+1} \right] &= \frac{d}{dx} [(x+1)(\ln(x+1) - \ln(x))] \\
 &= \ln(x+1) - \ln(x) + (x+1) \left[\frac{1}{x+1} - \frac{1}{x} \right] \\
 &= \ln(x+1) - \ln(x) - \frac{1}{x} \\
 &= \frac{1}{\eta} - \frac{1}{x} < 0, \quad \text{by (2.7).}
 \end{aligned}$$

Hence $(1 + \frac{1}{x})^{x+1}$ is a decreasing function of the variable x .

Example 2.5 The mean value theorem can be used in establishing the formula

$$\int_0^b x^\alpha dx = \frac{b^{\alpha+1}}{\alpha+1} \tag{2.8}$$

for $\alpha \geq 0$ and $b > 0$.

To establish (2.8), define a function f by $f(t) = \frac{t^{\alpha+1}}{\alpha+1}$. Then by the mean value theorem, there exists an $\eta \in]k-1, k[$ for every positive integer k such that

$$\frac{k^{\alpha+1}}{\alpha+1} - \frac{(k-1)^{\alpha+1}}{\alpha+1} = \eta^\alpha. \tag{2.9}$$

Since $\eta \in]k-1, k[$, we see that

$$(k-1)^\alpha < \eta^\alpha < k^\alpha.$$

Using (2.9) in the above inequalities, we get

$$(k-1)^\alpha < \frac{k^{\alpha+1}}{\alpha+1} - \frac{(k-1)^{\alpha+1}}{\alpha+1} < k^\alpha.$$

Summing over the index k from 1 to n , we obtain

$$\sum_{k=1}^n (k-1)^\alpha < \frac{n^{\alpha+1}}{\alpha+1} < \sum_{k=1}^n k^\alpha. \quad (2.10)$$

From inequalities (2.10) one can deduce

$$\frac{1}{\alpha+1} < \left(\frac{1^\alpha + 2^\alpha + \cdots + n^\alpha}{n^{\alpha+1}} \right) < \frac{1}{\alpha+1} + \frac{1}{n}.$$

Letting $n \rightarrow \infty$, we obtain

$$\lim_{n \rightarrow \infty} \left(\frac{1^\alpha + 2^\alpha + \cdots + n^\alpha}{n^{\alpha+1}} \right) = \frac{1}{\alpha+1}.$$

The definition of definite integral implies that

$$\int_0^b x^\alpha dx = \lim_{n \rightarrow \infty} \left(\frac{1^\alpha + 2^\alpha + \cdots + n^\alpha}{n^{\alpha+1}} \right) b^{\alpha+1}.$$

Thus we have

$$\int_0^b x^\alpha dx = \frac{b^{\alpha+1}}{\alpha+1}.$$

This approach has some advantages over the traditional approach found in many calculus textbooks.

Example 2.6 Let f be a function defined on $]a, b[$, and suppose $f'(c)$ exists for some $c \in]a, b[$. Let g be differentiable on an interval containing $f(c+h)$ for h sufficiently small, and suppose g' is continuous at $f(c)$. Then $g \circ f$ is differentiable at c and

$$(g \circ f)'(c) = g'(f(c))f'(c).$$

Since g is differentiable, by Lagrange's mean value theorem, we get

$$g(f(h+c)) - g(f(c)) = g'(\theta) [f(c+h) - f(c)]$$

for some θ strictly between $f(c+h)$ and $f(c)$. Now f is differentiable at c , so

$$\lim_{h \rightarrow 0} \frac{f(c+h) - f(c)}{h} = f'(c).$$

As f is continuous, $\lim_{h \rightarrow 0} f(c+h) = f(c)$ and thus $\lim_{h \rightarrow 0} \theta = f(c)$ (since θ is strictly between $f(c+h)$ and $f(c)$). Using the continuity of g' at $f(c)$, we have

$$\lim_{h \rightarrow 0} g'(\theta) = g' \left(\lim_{h \rightarrow 0} \theta \right) = g'(f(c)).$$

Hence

$$\begin{aligned} g'(f(c)) f'(c) &= \lim_{h \rightarrow 0} g'(\theta) \lim_{h \rightarrow 0} \frac{f(c+h) - f(c)}{h} \\ &= \lim_{h \rightarrow 0} g'(\theta) \left[\frac{f(c+h) - f(c)}{h} \right] \\ &= \lim_{h \rightarrow 0} \left[\frac{g(f(c+h)) - g(f(c))}{h} \right] \\ &= (g \circ f)'(c). \end{aligned}$$

Example 2.7 The mean value theorem can also be used in introducing an infinite family of means, known as Stolarsky's mean (see Stolarsky (1975)).

Define $f(x) = x^\alpha$, where α is a real parameter. We apply the mean value theorem to f on the interval $[x, y]$. There exists a point η with $x < \eta < y$ (which depends on the x, y and α) such that

$$f'(\eta_\alpha(x, y)) = \frac{f(x) - f(y)}{x - y}$$

which is

$$\eta_\alpha(x, y) = \left(\frac{x^\alpha - y^\alpha}{\alpha(x - y)} \right)^{\frac{1}{\alpha-1}}.$$

Note that we have used $\eta_\alpha(x, y)$ instead of η to emphasize the dependence of η on x, y and α . From this, one obtains an infinite family of means by

varying the parameter α . These means are known as Stolarsky's means. For instance, if $\alpha = -1$, then one gets the geometric mean

$$\eta_{-1}(x, y) = \sqrt{xy};$$

if $\alpha = 2$, then one gets the arithmetic mean

$$\eta_2(x, y) = \frac{x+y}{2};$$

if $\alpha \rightarrow 0$, then one gets the logarithmic mean

$$\lim_{\alpha \rightarrow 0} \eta_\alpha(x, y) = \frac{x-y}{\ln x - \ln y};$$

if $\alpha \rightarrow 1$, then one gets the identric mean

$$\lim_{\alpha \rightarrow 1} \eta_\alpha(x, y) = \frac{1}{e} \left[\frac{y^y}{x^x} \right]^{\frac{1}{y-x}}.$$

In the last example, we have seen that one can construct an infinite class of means given two positive real numbers x and y using the mean value theorem. It is easy to extend the definitions of the arithmetic and geometric means to n positive real numbers. It is obvious that the arithmetic mean of n positive numbers is

$$A(x_1, x_2, \dots, x_n) = \frac{x_1 + x_2 + \dots + x_n}{n}$$

whereas the geometric mean is

$$G(x_1, x_2, \dots, x_n) = \sqrt[n]{x_1 x_2 \cdots x_n}.$$

However, it is not so obvious to find an appropriate formula for the logarithmic mean of n positive numbers. In the remaining portion of this section we discuss how one can extend the definition of the logarithmic mean in the case of more than two positive real numbers. Recall that the logarithmic mean of positive real numbers x and y is given by

$$L(x, y) = \begin{cases} \frac{x-y}{\ln x - \ln y} & \text{if } x \neq y \\ x & \text{if } x = y. \end{cases}$$

Given three positive real numbers x , y and z one can construct an infinite class of means by using a different quotient which approximates a second

derivative. Using the mean value theorem for divided difference (see Theorem 2.10), we have

$$\frac{f(x)}{(x-y)(x-z)} + \frac{f(y)}{(y-z)(y-x)} + \frac{f(z)}{(z-x)(z-y)} = \frac{1}{2} f''(\eta)$$

where $\min\{x, y, z\} < \eta < \max\{x, y, z\}$. As in the previous example, we let $f(t) = t^\alpha$ to obtain

$$\eta_\alpha(x, y, z) = \left[\frac{2}{\alpha(\alpha-1)} \left\{ \frac{z^\alpha(y-x) + y^\alpha(x-z) + x^\alpha(z-y)}{(z-y)(z-x)(y-x)} \right\} \right]^{\frac{1}{\alpha-2}}.$$

If we put $\alpha = 3$, then we get

$$\eta_3(x, y, z) = \frac{x+y+z}{3};$$

with $\alpha = -1$, we get

$$\eta_{-1}(x, y, z) = \sqrt[3]{xyz}.$$

Two generalizations of the logarithmic mean can be constructed by considering the limiting cases of $\alpha = 0$ and $\alpha = 1$. For instance

$$\lim_{\alpha \rightarrow 0} \eta_\alpha(x, y, z) = \sqrt{\frac{(z-y)(z-x)(y-x)}{2 \left[x \ln\left(\frac{z}{y}\right) + y \ln\left(\frac{x}{z}\right) + z \ln\left(\frac{y}{x}\right) \right]}},$$

$$\lim_{\alpha \rightarrow 1} \eta_\alpha(x, y, z) = \frac{(z-y)(z-x)(y-x)}{2 \left[zy \ln\left(\frac{z}{y}\right) + xz \ln\left(\frac{x}{z}\right) + xy \ln\left(\frac{y}{x}\right) \right]}.$$

A generalization of the identric mean can be obtained by considering the limiting case when $\alpha = 2$. For example,

$$\lim_{\alpha \rightarrow 2} \eta_\alpha(x, y, z)$$

$$= \exp \left(-\frac{3}{2} + \frac{z^2 \ln z}{(z-x)(z-y)} + \frac{y^2 \ln y}{(y-x)(y-z)} + \frac{x^2 \ln x}{(x-y)(x-z)} \right).$$

Besides these generalizations, one can also generalize the logarithmic mean by examining the appropriate integral representation of the function

$L(x, y)$. It can be checked that

$$L(x, y) = \int_0^1 \left(\frac{x}{y}\right)^t y dt.$$

In view of this integral, one can define the logarithmic mean between three positive real numbers as

$$L(x, y, z) = \int_0^1 \int_0^1 \left(\frac{x}{z}\right)^t \left(\frac{y}{z}\right)^s z dt ds.$$

Evaluating the above integral we obtain the following explicit form of L as

$$L(x, y, z) = \frac{(z - x)(z - y)}{z(\ln z - \ln x)(\ln z - \ln y)}.$$

The above can be generalized in the case of n positive real numbers x_1, x_2, \dots, x_n as

$$L(x_1, x_2, \dots, x_n)$$

$$= \int_0^1 \int_0^1 \cdots \int_0^1 \left(\frac{x_1}{x_n}\right)^{t_1} \left(\frac{x_2}{x_n}\right)^{t_2} \cdots \left(\frac{x_{n-1}}{x_n}\right)^{t_{n-1}} x_n dt_1 dt_2 \cdots dt_{n-1}.$$

It should also be noted that the function $L(x, y)$ can also be represented by the following integral. That is

$$L(x, y) = \left[\int_0^1 \frac{dt}{tx + (1-t)y} \right]^{-1}.$$

In view of the above integral representation of the logarithmic mean, we have the following extension

$$L(x_1, x_2, \dots, x_n) = \left[(n-1)! \int_{\Gamma_n} \left(\sum_{i=1}^n t_i x_i \right)^{-1} dt \right]^{-1},$$

where $\Gamma_n = \left\{ (t_1, t_2, \dots, t_n) \mid t_i \geq 0, \sum_{i=1}^n t_i \leq 1 \right\}$, $t_n = 1 - \sum_{i=1}^{n-1} t_i$, and $dt = dt_1 dt_2 \cdots dt_{n-1}$.

2.3 Associated Functional Equations

In this section, we illustrate a functional equation that arises from the mean value theorem and then we present a systematic study of this functional equation and its various generalizations. These functional equations characterize polynomials of various degrees. To this end we need some notation which we introduce at this point:

Definition 2.1 For distinct real numbers x_1, x_2, \dots, x_n , the *divided difference* of the function $f : \mathbb{R} \rightarrow \mathbb{R}$ is defined as

$$f[x_1] = f(x_1)$$

and

$$f[x_1, x_2, \dots, x_n] = \frac{f[x_1, x_2, \dots, x_{n-1}] - f[x_2, x_3, \dots, x_n]}{x_1 - x_n}$$

for all $n \geq 2$.

It is easy to see that

$$f[x_1, x_2] = \frac{f(x_1) - f(x_2)}{x_1 - x_2}$$

and

$$f[x_1, x_2, x_3] = \frac{(x_3 - x_2)f(x_1) + (x_1 - x_3)f(x_2) + (x_2 - x_1)f(x_3)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}.$$

In view of the definition of the divided difference, equation (2.1) in the mean value theorem takes the form

$$f[x_1, x_2] = f'(\eta(x_1, x_2)). \quad (2.11)$$

Obviously η depends on x_1 and x_2 and one may ask for what f the mean value η depends on x_1 and x_2 in a given manner. From this point of view, equation (2.11) appears as a functional equation with unknown function f and given η .

The following theorem was established by Aczél (1963) and also independently by Haruki (1979). The proof of the following theorem is based on the proof which appeared in Aczél (1985). This theorem is related to equation (2.11).

Theorem 2.3 *The functions $f, h : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$f[x, y] = h(x + y), \quad x \neq y, \quad (2.12)$$

if and only if

$$f(x) = ax^2 + bx + c \quad \text{and} \quad h(x) = ax + b$$

where a, b, c are arbitrary real constants.

Proof: Equation (2.12), using the definition of the divided difference of f , can be rewritten as

$$f(x) - f(y) = (x - y) h(x + y) \quad \text{for } x \neq y \quad (2.13)$$

which is also true for $x = y$. If f satisfies equation (2.13), so does $f + b$, where b is an arbitrary constant. Therefore we may assume without loss of generality $f(0) = 0$. Putting $y = 0$ in equation (2.13), we see that

$$f(x) = x h(x). \quad (2.14)$$

Hence by equation (2.14), equation (2.13) transforms into

$$x h(x) - y h(y) = (x - y) h(x + y). \quad (2.15)$$

Again if h satisfies equation (2.15) so also $h + c$, where c is an arbitrary constant. So we may suppose $h(0) = 0$. Therefore, letting $x = -y$ in equation (2.15), we obtain

$$-y h(-y) = y h(y) \quad (2.16)$$

that is h is an odd function. Taking this into consideration and replacing y by $-y$ in equation (2.15), we get

$$x h(x) - y h(y) = (x + y) h(x - y). \quad (2.17)$$

Comparing equation (2.17) with equation (2.15), we have

$$(x - y) h(x + y) = (x + y) h(x - y) \quad (2.18)$$

and substituting

$$u = x + y \quad \text{and} \quad v = x - y \quad (2.19)$$

in equation (2.18), we obtain

$$v h(u) = u h(v) \quad (2.20)$$

for all $u, v \in \mathbb{R}$. Thus

$$h(u) = au. \quad (2.21)$$

If we do not assume $h(0) = 0$, then we have in general

$$h(u) = au + b.$$

By equation (2.14) this gives $f(x) = x(ax + b)$ and, if we do not assume $f(0) = 0$, then

$$f(x) = ax^2 + bx + c.$$

So we have indeed proved that all solutions of equation (2.12) are of the form

$$\begin{aligned} f(x) &= ax^2 + bx + c \\ h(x) &= ax + b, \end{aligned}$$

where a, b, c are arbitrary constants, as asserted. The converse of this theorem is straightforward and the proof is now complete.

The following corollary follows from Theorem 2.3.

Corollary 2.1 *The function $f : \mathbb{R} \rightarrow \mathbb{R}$ satisfies the functional equation*

$$f(x) - f(y) = (x - y) f' \left(\frac{x+y}{2} \right), \quad x \neq y,$$

if and only if

$$f(x) = ax^2 + bx + c$$

where a, b, c are arbitrary real constants.

Theorem 2.4 *If the quadratic polynomial $f(x) = ax^2 + bx + c$ with $a \neq 0$, is a solution of the functional equation*

$$f(x + h) - f(x) = h f'(x + \theta h) \quad (0 < \theta < 1) \quad (2.22)$$

assumed for all $x \in \mathbb{R}$ and $h \in \mathbb{R} \setminus \{0\}$, then $\theta = \frac{1}{2}$. Conversely, if a function f satisfies the above functional differential equation with $\theta = \frac{1}{2}$, then the only solution is a polynomial of degree at most two.

Proof: Suppose the polynomial

$$f(x) = ax^2 + bx + c \quad (2.23)$$

is a solution of equation (2.22). Then inserting equation (2.23) into (2.22), we have

$$a(x+h)^2 + b(x+h) + c - ax^2 - bx - c = h(2a(x+\theta h) + b)$$

that is

$$ah^2(1-2\theta) = 0.$$

Since a and h are nonzero, we have

$$\theta = \frac{1}{2}.$$

This proves the *if* part of the theorem. Next, we prove the converse of the theorem. Letting $\theta = \frac{1}{2}$ and $h = y - x$ in equation (2.22), we see that

$$f(x) - f(y) = (x-y)f'\left(\frac{x+y}{2}\right), \quad x \neq y.$$

Thus, by Corollary 1, f is a polynomial of degree at most two and the proof of theorem is now complete.

Let s and t be given real numbers. Then all differentiable functions f on the real line which satisfy

$$f[x, y] = f'(sx + ty) \quad (2.24)$$

for all real numbers x, y , with $x \neq y$ are of the form

$$f(x) = \begin{cases} ax^2 + bx + c & \text{if } s = t = \frac{1}{2} \\ bx + c & \text{otherwise,} \end{cases}$$

where a, b, c are arbitrary constants (see E3338 (1991) and Rudin (1989)). This result was shown independently by Baker, Jacobson and Sahoo, and Falkowitz in 1991.

In a recent paper, Kannappan, Sahoo and Jacobson (1995) established the following theorem. We present the proof of the following theorem based on their proof.

Theorem 2.5 Let s and t be the real parameters. Functions $f, g, h : \mathbb{R} \rightarrow \mathbb{R}$ satisfy

$$\frac{f(x) - g(y)}{x - y} = h(sx + ty) \quad (2.25)$$

for all $x, y \in \mathbb{R}$, $x \neq y$ if and only if

$$f(x) = \begin{cases} ax + b & \text{if } s = 0 = t \\ ax + b & \text{if } s = 0, t \neq 0 \\ ax + b & \text{if } s \neq 0, t = 0 \\ \alpha tx^2 + ax + b & \text{if } s = t \neq 0 \\ \frac{A(tx)}{t} + b, & \text{if } s = -t \neq 0 \\ \beta x + b & \text{if } s^2 \neq t^2 \end{cases} \quad (2.26)$$

$$g(y) = \begin{cases} ay + b & \text{if } s = 0 = t \\ ay + b & \text{if } s = 0, t \neq 0 \\ ay + b & \text{if } s \neq 0, t = 0 \\ \alpha ty^2 + ay + b & \text{if } s = t \neq 0 \\ \frac{A(ty)}{t} + c, & \text{if } s = -t \neq 0 \\ \beta y + b & \text{if } s^2 \neq t^2 \end{cases} \quad (2.27)$$

$$h(y) = \begin{cases} \text{arbitrary with } h(0) = a & \text{if } s = 0 = t \\ a & \text{if } s = 0, t \neq 0 \\ a & \text{if } s \neq 0, t = 0 \\ \alpha y + a & \text{if } s = t \neq 0 \\ \frac{A(y)}{y} + \frac{(c-b)t}{y}, & \text{if } s = -t \neq 0, y \neq 0 \\ \beta & \text{if } s^2 \neq t^2, \end{cases} \quad (2.28)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function and a, b, c, α, β are arbitrary real constants.

Proof: To prove the theorem, we consider several cases depending on parameters s and t .

Case 1. Suppose $s = 0 = t$. Then equation (2.25) yields

$$\frac{f(x) - g(y)}{x - y} = h(0)$$

which is

$$f(x) - ax = g(y) - ay,$$

where $a = h(0)$. From the above, we obtain

$$f(x) = ax + b \quad \text{and} \quad g(y) = ay + b, \quad (2.29)$$

where b is an arbitrary constant. Letting equation (2.29) into equation (2.25), we see that h is an arbitrary function with $a = h(0)$. Thus we obtain the solution as asserted in theorem for the case $s = 0 = t$.

Case 2. Suppose $s = 0$ and $t \neq 0$. (The case $s \neq 0$ and $t = 0$ can be handled in a similar manner.) Then from equation (2.25), we get

$$\frac{f(x) - g(y)}{x - y} = h(ty). \quad (2.30)$$

Putting $y = 0$ in equation (2.30), we see that

$$f(x) = ax + b, \quad x \neq 0 \quad (2.31)$$

where $a = h(0)$ and $b = g(0)$. Letting equation (2.31) into equation (2.30), we obtain

$$ax + b - g(y) = (x - y)h(ty) \quad (2.32)$$

for all $x \neq y$ and $x \neq 0$. Equating the coefficients of x and the constant terms in equation (2.32), we get

$$h(ty) = a \quad \text{and} \quad g(y) = h(ty)y + b = ay + b \quad (2.33)$$

for all $y \in \mathbb{R}$. Letting $x = 0$ in equation (2.30) and using equation (2.33), we see that $f(0) = b$. Thus equation (2.31) holds for all x in \mathbb{R} . Hence from equation (2.31) and equation (2.33), we get the solution of (2.25) for this case as asserted in Theorem 2.5.

Case 3. Suppose $s \neq 0 \neq t$. Letting $x = 0$ in equation (2.25), we get

$$g(y) = yh(ty) + b \quad (2.34)$$

for all real $y \neq 0$ (where $b = f(0)$). Similarly, letting $y = 0$ in equation (2.25), we get

$$f(x) = xh(sx) + c \quad (2.35)$$

for all $x \neq 0$ (where $c = g(0)$). Inserting equation (2.34) and equation (2.35) into equation (2.25) and simplifying, we obtain

$$xh(sx) - yh(ty) + c - b = (x - y)h(sx + ty) \quad (2.36)$$

for all real nonzero x and y with $x \neq y$.

Replacing x by $\frac{x}{s}$ and y by $\frac{y}{t}$ in equation (2.36), we get

$$\frac{x}{s} h(x) - \frac{y}{t} h(y) + c - b = \left(\frac{x}{s} - \frac{y}{t} \right) h(x+y) \quad (2.37)$$

for all real nonzero x and y with $tx \neq sy$.

Subcase 3.1. Suppose $s = t$. Hence equation (2.37) yields

$$x h(x) - y h(y) = (b - c) t + (x - y) h(x+y). \quad (2.38)$$

Interchanging x with y in equation (2.38) and adding the resulting equation to (2.38) we get $b = c$. Thus equation (2.38) reduces to

$$x h(x) - y h(y) = (x - y) h(x+y) \quad (2.39)$$

for all real nonzero x and y with $x \neq y$. Replacing y with $-y$ in equation (2.39), we obtain

$$x h(x) + y h(-y) = (x + y) h(x - y) \quad (2.40)$$

for all real nonzero x and y with $x + y \neq 0$. Letting $y = -x$ in equation (2.39), we see that

$$x h(x) + x h(-x) = 2x h(0). \quad (2.41)$$

Subtracting equation (2.39) from equation (2.40) and using equation (2.41), we get

$$2y h(0) = (x + y) h(x - y) - (x - y) h(x + y) \quad (2.42)$$

for all real nonzero x, y with $x + y$ and $x - y \neq 0$. Writing

$$u = x + y \quad \text{and} \quad v = x - y \quad (2.43)$$

in equation (2.42), we see that

$$(u - v) h(0) = u h(v) - v h(u)$$

which is

$$v [h(u) - h(0)] = u [h(v) - h(0)], \quad (2.44)$$

for all real nonzero $u, v, u - v$ and $u + v$. Thus

$$h(u) = \alpha u + a \quad (2.45)$$

for all real nonzero u in \mathbb{R} (where $a = h(0)$). Notice that equation (2.45) also holds for $u = 0$. Using equation (2.45) in equation (2.25), we get

$$f(x) - g(y) = (x - y)(\alpha t x + \alpha t y + a)$$

for all $x \neq y$. Thus, we obtain the asserted solution

$$f(x) = g(x) = \alpha t x^2 + a x + b \quad \text{and} \quad h(y) = \alpha y + a, \quad (2.46)$$

where α, a and b are arbitrary constants.

Subcase 3.2. Suppose $s = -t$. Then equation (2.37) yields

$$x h(x) + y h(y) + (b - c)t = (x + y) h(x + y) \quad (2.47)$$

for all real nonzero x and y with $x \neq y$. Define

$$A(x) = \begin{cases} x h(x) + (b - c)t & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases} \quad (2.48)$$

Then by equation (2.48), equation (2.47) reduces to

$$A(x) + A(y) = A(x + y) \quad (2.49)$$

for all real nonzero x, y and $x + y$. Next we show that A in equation (2.49) is additive on the set of reals. In order for A to be additive it must satisfy

$$\begin{aligned} A(x) + A(-x) &= A(0) = 0 \\ \text{or} \\ x h(x) - x h(-x) + 2(b - c)t &= 0. \end{aligned} \quad (2.50)$$

Interchanging y with $-y$ in equation (2.47), we get

$$x h(x) - y h(-y) + (b - c)t = (x - y) h(x - y). \quad (2.51)$$

Subtracting equation (2.51) from equation (2.47), we get

$$y h(y) + y h(-y) = (x + y) h(x + y) - (x - y) h(x - y). \quad (2.52)$$

Thus, using equation (2.48), we get

$$A(y) - A(-y) = A(x + y) - A(x - y) \quad (2.53)$$

for all real nonzero $x, y, x + y$ and $x - y$. Replacing x by $-x$ in equation (2.53), we obtain

$$A(y) - A(-y) = A(-x + y) - A(-x - y). \quad (2.54)$$

From equation (2.53) and equation (2.54), we get

$$A(x + y) + A(-(x + y)) = A(x - y) + A(-(x - y)). \quad (2.55)$$

Letting $u = x + y$ and $v = x - y$ in equation (2.55), we see that

$$A(u) + A(-u) = A(v) + A(-v) \quad (2.56)$$

for all real nonzero $u, v, u - v$ and $u + v$. Thus

$$A(u) + A(-u) = \gamma \quad (2.57)$$

for all real nonzero u (where γ is a constant). Using equation (2.48), we see from equation (2.57) that

$$x h(x) - x h(-x) + 2(b - c)t = \gamma, \quad (2.58)$$

for all real nonzero x . From equation (2.25) with $s = -t$, we get

$$f(x) - g(y) = (x - y) h(-(x - y)t). \quad (2.59)$$

Interchanging x with y , we get

$$f(y) - g(x) = -(x - y) h((x - y)t). \quad (2.60)$$

Adding equation (2.59) to equation (2.60) and using equation (2.58), we get

$$\begin{aligned} & f(x) - g(x) + f(y) - g(y) \\ &= -(x - y) h((x - y)t) + (x - y) h(-(x - y)t) \\ &= -\frac{\gamma}{t} + 2(b - c). \end{aligned} \quad (2.61)$$

Using equation (2.34) and equation (2.48), we obtain

$$A(tx) = t [g(x) - c] \quad (x \neq 0). \quad (2.62)$$

Similarly, using equation (2.35) and equation (2.48), we get

$$A(-tx) = -t [f(x) - b] \quad (x \neq 0). \quad (2.63)$$

So from equation (2.62) and equation (2.63), we see that

$$\begin{aligned} f(x) - g(x) &= -\frac{A(-tx) + A(tx)}{t} + b - c \\ &= -\frac{\gamma}{t} + b - c. \end{aligned} \quad (2.64)$$

Hence from above, we get

$$f(x) - g(x) + f(y) - g(y) = -2\frac{\gamma}{t} + 2(b - c). \quad (2.65)$$

Comparing equation (2.61) with equation (2.65), we get $\gamma = 0$. Thus equation (2.57) yields

$$A(x) + A(-x) = 0, \quad (2.66)$$

for all real nonzero x . Evidently the above also holds for $x = 0$. Hence A is an additive function on the set of reals. From equation (2.48), equation (2.34) and equation (2.35), we obtain

$$\left. \begin{array}{lcl} f(x) &= & \frac{A(tx)}{t} + b \\ g(y) &= & \frac{A(ty)}{t} + c \\ h(y) &= & \frac{A(y)}{y} + \frac{(c-b)t}{y}, \end{array} \right\} \quad (2.67)$$

where b and c are arbitrary constants.

Subcase 3.3. Suppose $s^2 \neq t^2$, that is $s \neq \pm t$. Interchanging x with y in equation (2.37), we get

$$\frac{y}{s} h(y) - \frac{x}{t} h(x) + c - b = \left(\frac{y}{s} - \frac{x}{t} \right) h(x+y) \quad (2.68)$$

for all nonzero x and y with $ty \neq sx$. Subtracting equation (2.68) from equation (2.37) and using $s+t \neq 0$, we get

$$x h(x) - y h(y) = (x-y) h(x+y), \quad (2.69)$$

which is the same as equation (2.39). Thus

$$h(x) = \alpha x + b, \quad (2.70)$$

where α and b are arbitrary constants. Letting equation (2.70) into equation (2.68) and simplifying the resulting expression, we get

$$\alpha xy \left(\frac{1}{s} - \frac{1}{t} \right) = b - c \quad (2.71)$$

for all nonzero x and y with $tx \neq sy$ and $sx \neq ty$. Since $s \neq t$, we see that $\alpha = 0$ and $b = c$. Hence equation (2.70) becomes

$$h(x) = b. \quad (2.72)$$

From equation (2.72), equation (2.34) and equation (2.35), we obtained the asserted form of f , g and h . This completes the proof of the theorem.

Remark 1. In case $g = f$, the Subcase 3.2 simplifies to a great extent. If $g = f$, then the left side of (2.25) for $s = -t$ is symmetric in x and y . Thus using this symmetry one can conclude that h is an even function. The evenness of h implies that A in equation (2.49) is additive.

Remark 2. In the Subcase 3.1, $h(y)$ is undefined at $y = 0$.

Remark 3. We have seen in Chapter 1 that the Cauchy functional equation $A(x + y) = A(x) + A(y)$ has discontinuous solutions in addition to the continuous solutions of the form $A(x) = ax$, where a is an arbitrary real constant. Since an additive function appears in the solution of (2.25) for the subcase $s = -t$, it follows that (2.25) has discontinuous solutions. However, all continuous solutions of (2.25) are polynomials of low degree.

The following corollary is obvious from the above theorem.

Corollary 2.2 *The functions $\phi, f : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation (2.11) for all $x, y \in \mathbb{R}$ with $x \neq y$ if and only if*

$$\begin{aligned} f(x) &= \begin{cases} ax + b & \text{if } s = 0 = t \\ ax + c & \text{if } s = 0, t \neq 0 \\ ax + c & \text{if } s \neq 0, t = 0 \\ \alpha tx^2 + ax + b & \text{if } s = t \neq 0 \\ \frac{A(tx)}{t} + b, & \text{if } s = -t \neq 0 \\ \beta x + b & \text{if } s^2 \neq t^2 \\ \text{arbitrary} & \text{if } s = 0 = t \\ a & \text{if } s = 0, t \neq 0 \\ a & \text{if } s \neq 0, t = 0 \end{cases} \\ \phi(y) &= \begin{cases} \alpha y + a & \text{if } s = t \neq 0 \\ \frac{A(y)}{y}, & \text{if } s = -t \neq 0, y \neq 0 \\ \beta & \text{if } s^2 \neq t^2, \end{cases} \end{aligned}$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function and a, b, c, α, β are arbitrary real constants.

The following corollary addresses a recreational problem posed by Walter Rudin in 1989.

Corollary 2.3 *The function $f : \mathbb{R} \rightarrow \mathbb{R}$ satisfies the equation*

$$f'(sx + ty) = \frac{f(y) - f(x)}{y - x}$$

for all $x, y \in \mathbb{R}$ with $x \neq y$ if and only if

$$f(x) = \begin{cases} ax^2 + bx + c & \text{if } s = \frac{1}{2} = t \\ bx + c & \text{otherwise,} \end{cases}$$

where a, b and c are arbitrary real constants.

Polynomials are basic objects in mathematics. In many applications functions are used for modeling of real world problems. These functions, in turn, provided they are sufficiently smooth, can be approximated by polynomials, in some range and within some accuracy. Polynomials play a central role in mathematics, for instance, in analysis, number theory, approximation theory, and numerical analysis.

Notice that the Theorem 2.5 and Theorem 2.3 characterize low degree polynomials. Originally the equation (2.12) appeared in the form

$$f(x) - f(y) = (x - y)h(x + y)$$

and thus it was not clear what the generalization of it would be for higher order polynomials. Ideas came from the notions of divided difference. Bailey (1992) generalized Aczél's and Haruki's result and established the following theorem.

Theorem 2.6 *If f is a differentiable function satisfying the functional equation*

$$f[x, y, z] = h(x + y + z), \quad (2.73)$$

then f is a polynomial of degree at most three.

Proof: Using the definition of divided difference, from equation (2.73), one obtains

$$\begin{aligned} f(x)(y-z) + f(y)(z-x) + f(z)(x-y) \\ = (x-y)(y-z)(x-z)h(x+y+z). \end{aligned} \quad (2.74)$$

If f satisfy equation (2.74) so also $f + d$, where d is an arbitrary constant. Therefore, we may assume without loss of generality $f(0) = 0$. Under this assumption we set $z = 0$ in equation (2.74) and obtain

$$yf(x) - xf(y) = xy(x-y)h(x+y). \quad (2.75)$$

Rewriting equation (2.75), we have

$$\frac{f(x)}{x} - \frac{f(y)}{y} = (x-y)h(x+y). \quad (2.76)$$

Now under the assumption that f is differentiable, h is continuous and thus, if we allow y to approach 0 on each side of equation (2.76), we obtain

$$f'(0) - \frac{f(x)}{x} = -xh(x). \quad (2.77)$$

Therefore, if we define

$$q(x) = \begin{cases} \frac{f(x)}{x} & \text{if } x \neq 0 \\ f'(0) & \text{if } x = 0 \end{cases}$$

we have $f(x) = xq(x)$ for all x and

$$q(y) - q(x) = (y-x)h(x+y).$$

By Theorem 2.3, we obtain

$$q(x) = ax^2 + bx + c$$

so that

$$f(x) = ax^3 + bx^2 + cx.$$

Removing the assumption that $f(0) = 0$ we have

$$f(x) = ax^3 + bx^2 + cx + d,$$

as asserted in the theorem. The proof is now complete.

Without being aware of the result of Crstici and Neagu (1987), in 1992 Bailey posed the question whether every continuous (or differentiable) f satisfying the functional equation

$$f[x_1, x_2, \dots, x_n] = g(x_1 + x_2 + \dots + x_n) \quad (2.78)$$

is a polynomial of degree at most n . Using some elementary techniques Kannappan and Sahoo (1995) have solved Bailey's problem. First, we solve Bailey's problem for $n = 3$ and then in the next theorem we present the solution of (2.78).

Theorem 2.7 . *Let f satisfy the functional equation*

$$f[x_1, x_2, x_3] = g(x_1 + x_2 + x_3), \quad (2.79)$$

for all $x_1, x_2, x_3 \in \mathbb{R}$ with $x_1 \neq x_2$, $x_2 \neq x_3$ and $x_3 \neq x_1$. Then f is a polynomial of degree at most three and g is linear.

Proof: If $f(x)$ is a solution of (2.79) so also $f(x) + a_0 + a_1x$. Hence we may assume without loss of generality that

$$f(0) = 0 \quad (2.80)$$

and

$$f(\alpha) = 0 \quad (2.81)$$

for some $\alpha \neq 0$ in \mathbb{R} . Note that there are many choices for such an α . First substitute $(x, 0, \alpha)$ for (x_1, x_2, x_3) in (2.79) to get

$$f(x) = -x(\alpha - x)g(x + \alpha) \quad (2.82)$$

(after using (2.80) and (2.81)) for $x \neq 0, \alpha$.

Next, we substitute $(x, 0, y)$ for (x_1, x_2, x_3) in (2.79) to get

$$\frac{f(x)}{x(x-y)} - \frac{f(y)}{y(x-y)} = g(x+y) \quad (2.83)$$

for all $x, y \neq 0$ and $x \neq y$. Define

$$q(x) = \frac{f(x)}{x} \quad (2.84)$$

for $x \in \mathbb{R} \setminus \{0\}$. Then (2.83) reduces to

$$q(x) - q(y) = (x - y) g(x + y) \quad (2.85)$$

for all $x, y \in \mathbb{R} \setminus \{0\}$ with $x \neq y$. Note that (2.85) is valid even for $x = y$.

Now we consider the equation

$$q(x) - q(y) = (x - y) g(x + y)$$

for all $x, y \in \mathbb{R} \setminus \{0\}$. Put $y = -x$ in (2.85) to get

$$q(x) - q(-x) = 2x g(0) \quad (2.86)$$

for all $x \neq 0$. Next, replace y by $-y$ in (2.85) to get

$$q(x) - q(-y) = (x + y) g(x - y) \quad (2.87)$$

with $x, y \in \mathbb{R} \setminus \{0\}$ with $x + y \neq 0$. Again (2.87) holds if $x + y = 0$. Thus we conclude that (2.87) holds for $x, y \in \mathbb{R} \setminus \{0\}$.

Subtract (2.85) from (2.87) and use (2.86) to get

$$(x + y) [g(x - y) - g(0)] = (x - y) [g(x + y) - g(0)] \quad (2.88)$$

for all $x, y \in \mathbb{R} \setminus \{0\}$. Fix a nonzero u in \mathbb{R} . Choose a $v \in \mathbb{R}$ such that $\frac{u+v}{2} \neq 0$ and $\frac{u-v}{2} \neq 0$. There are plenty of choices for such v . Let

$$x = \frac{u+v}{2} \quad \text{and} \quad y = \frac{u-v}{2}$$

so that

$$u = x + y \quad \text{and} \quad v = x - y. \quad (2.89)$$

Letting (2.89) into (2.88), we get

$$u [g(v) - g(0)] = v [g(u) - g(0)] \quad (2.90)$$

for all $v \neq u, -u$. (Here note that v can be zero since $x = y$ is allowed.) Hence for fixed $u = u_1$, we get

$$g(v) = a_1 v + b_1 \quad (2.91)$$

for $v \in \mathbb{R} \setminus \{u_1, -u_1\}$. Again for $u = u_2$, we get

$$g(v) = a_2 v + b_2 \quad (2.92)$$

for all $v \in \mathbb{R} \setminus \{u_2, -u_2\}$. Since the sets $\{u_1, -u_1\}$ and $\{u_2, -u_2\}$ are disjoint, we get

$$g(v) = av + b \quad (2.93)$$

for all $v \in \mathbb{R}$. Now using (2.93) in (2.83), we have

$$\begin{aligned} f(x) &= (x^2 - x\alpha) g(x + \alpha) \\ &= (x^2 - x\alpha) [a(x + \alpha) + b] \\ &= ax^3 + bx^2 + cx \end{aligned}$$

where $c = -a\alpha^2 - b\alpha$. Removing the assumption that $f(0) = 0$, we get

$$f(x) = ax^3 + bx^2 + cx + d \quad (2.94)$$

for all $x \neq 0, \alpha$. By (2.80), (2.81) and (2.94), we conclude that f is a polynomial of degree at most three for all $x \in \mathbb{R}$. This proof is now complete.

Now we find the solution of (2.78) without any assumptions on the unknown function f and g . The following lemma is needed to solve Bailey's problem.

Lemma 2.5 *Let S be a finite subset of \mathbb{R} symmetric about zero (that is, $-S = S$) and let $f, g : \mathbb{R} \rightarrow \mathbb{R}$ be functions satisfying the functional equation*

$$f(x) - f(y) = (x - y) g(x + y) \quad (2.95)$$

for all $x, y \in \mathbb{R} \setminus S$. Then

$$f(x) = ax^2 + bx + c \quad \text{and} \quad g(y) = ay + b \quad (2.96)$$

for $x \in \mathbb{R} \setminus S$ and $y \in \mathbb{R}$, where a, b, c are some constants.

Proof: Replacing y by $-x$ in (2.95), we obtain

$$f(x) - f(-x) = 2xg(0), \quad \text{for } x \in \mathbb{R} \setminus S \quad (2.97)$$

Again, replacing y by $-y$ in (2.95), we get

$$f(x) - f(-y) = (x + y)g(x - y), \quad \text{for } x, y \in \mathbb{R} \setminus S$$

which after subtracting from (2.95) and using (2.97) gives

$$(x+y)(g(x-y)-g(0)) = (x-y)(g(x+y)-g(0)) \quad (2.98)$$

for all $x, y \in \mathbb{R} \setminus S$. Fix a nonzero $u \in \mathbb{R}$. Let $v \in \mathbb{R}$ such that $(u \pm v)/2 \notin S$ and put $x = (u+v)/2$ and $y = (u-v)/2$. Then $x+y = u$ and $x-y = v$ and use (2.98) to get

$$u(g(v)-g(0)) = v(g(u)-g(0)) \quad (2.99)$$

for all $v \in \mathbb{R} \setminus (2S \pm u)$, where $2S \pm u$ denotes the set

$$\{2s+u \mid s \in S\} \cup \{2s-u \mid s \in S\}.$$

For each fixed u , the equation (2.99) shows that g is linear in v , that is of the form $av+b$, except on the finite set $2S \pm u$. To conclude that g is linear on the reals, one has to note that, if one takes two suitable different values of u , which is now treated as a parameter, the exceptional sets involved are disjoint and so $g(v) = av+b$ for all real v with the same constants everywhere.

Substituting this for g in (2.95), we obtain

$$f(x) - ax^2 - bx = f(y) - ay^2 - by \quad (2.100)$$

for all $x, y \in \mathbb{R} \setminus S$. Choosing any $y \in \mathbb{R} \setminus S$ in (2.100) yields that $f(x) = ax^2 + bx + c$ for $x \in \mathbb{R} \setminus S$, for some constant c , which is the required form of f in (2.96). This completes the proof of the lemma.

The following theorem addresses a the problem posed by Bailey in 1992.

Theorem 2.8 *Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation (2.78) for distinct x_1, x_2, \dots, x_n , that is, for $x_i \neq x_j$ ($i \neq j$, $i, j = 1, 2, \dots, n$). Then f is a polynomial of degree at most n and g is linear, that is, a polynomial of first degree.*

Proof: It is easy to see that if f is a solution of equation (2.78), so also $f(x) + \sum_{k=0}^{n-2} a_k x^k$. So, we can assume that $f(0) = 0 = f(y_1) = \dots = f(y_{n-2})$ for y_1, y_2, \dots, y_{n-2} distinct and different from zero. Obviously there are plenty of choices for $0, y_1, \dots, y_{n-2}$. Putting in equation (2.78),

$(x, 0, y_1, \dots, y_{n-2})$ and $(x, 0, y, y_1, \dots, y_{n-3})$ for (x_1, x_2, \dots, x_n) , we get

$$f(x) = -x(y_1 - x) \cdots (y_{n-2} - x) g\left(x + \sum_{k=1}^{n-2} y_k\right) \quad (2.101)$$

and

$$\begin{aligned} & \frac{f(x)}{x(x-y)(y_1-x)\cdots(y_{n-3}-x)} - \frac{f(y)}{y(x-y)(y_1-y)\cdots(y_{n-3}-y)} \\ &= g\left(x+y+\sum_{k=1}^{n-3} y_k\right) \end{aligned}$$

respectively for $x \neq y$ and $x \neq 0, y, y_1, \dots, y_{n-2}$.

Now the above equation can be rewritten as

$$l(x) - l(y) = (x-y) g\left(x+y+\sum_{k=1}^{n-3} y_k\right), \quad (2.102)$$

where $l(x) = \frac{f(x)}{x(y_1-x)\cdots(y_{n-3}-x)}$ for $x, y \neq 0, y_1, \dots, y_{n-3}$. Then by Lemma 2.5 and the arbitrary choice of $x, y \neq 0, y_1, \dots, y_{n-3}$ we get that g is linear (and $l(x)$ is quadratic). Hence by (2.101), f is a polynomial of degree at most n . This proves the theorem.

Note that the above theorem is a straight forward generalization of the previous theorem. Schwaiger (1995) has also established this result independently.

The functional equation

$$f[x, y] = h(\eta(x, y)) \quad (2.103)$$

has been studied by taking η to be geometric mean and harmonic mean of x and y (see Aczél and Kuczma (1989)). Further, the functional equation (2.103) was treated in Kuczma (1991) assuming $\eta(x, y)$ to be a quasiarithmetic mean. A similar indepth study of the following functional equation

$$f[x_1, x_2, \dots, x_n] = h(\eta(x_1, x_2, \dots, x_n))$$

has not been done by taking $\eta(x_1, x_2, \dots, x_n)$ to be a quasiarithmetic mean.

2.4 The MVT for Divided Differences

In this section, we prove the mean value theorem for divided differences and then present some applications toward the study of means. We begin this section with an integral representation of divided differences. Some of the results of this section can be found in the books of Isaacson and Keller (1966) and Ostrowski (1973). In this section $f^{(n)}$ will denote the n^{th} derivative of a function f while f' will represent the first derivative of f .

Theorem 2.9 *Suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ has a continuous n^{th} derivative in the interval*

$$\min\{x_o, x_1, \dots, x_n\} \leq x \leq \max\{x_o, x_1, \dots, x_n\}.$$

If the points x_o, x_1, \dots, x_n are all distinct, then

$$\begin{aligned} & \int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} f^{(n)} \left(x_o + \sum_{k=1}^n t_k (x_k - x_{k-1}) \right) dt_n \\ &= f[x_o, x_1, \dots, x_n] \end{aligned} \tag{2.104}$$

where $n \geq 1$.

Proof: We prove this theorem by induction. If $n = 1$, the representation given in (2.104) reduces to

$$f[x_o, x_1] = \int_0^1 f'(t_1(x_1 - x_o) + x_o) dt_1.$$

First we show that the integral on the right side of the above equation is equal to the divided difference of f based on the two distinct points x_o and x_1 . Consider the integral

$$\int_0^1 f'(t_1(x_1 - x_o) + x_o) dt_1.$$

Since $x_1 \neq x_o$, introducing a new variable z for $t_1(x_1 - x_o) + x_o$, we get

$$dz = (x_1 - x_o) dt_1,$$

that is

$$dt_1 = \frac{dz}{x_1 - x_o}.$$

Since $t_1 = 0$, the new variable $z = x_o$ and similarly if $t_1 = 1$, then $z = x_1$. Hence, we have

$$\begin{aligned} \int_0^1 f'(t_1(x_1 - x_o) + x_o) dt_1 &= \int_{x_o}^{x_1} f'(z) \frac{dz}{x_1 - x_o} \\ &= \frac{\int_{x_o}^{x_1} f'(z) dz}{x_1 - x_o} \\ &= \frac{f(x_1) - f(x_o)}{x_1 - x_o} \\ &= f[x_o, x_1]. \end{aligned}$$

Next, assuming that the integral representation in (2.104) holds for $n - 1$, that is

$$\begin{aligned} &\int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-2}} f^{(n-1)} \left(x_o + \sum_{k=1}^{n-1} t_k (x_k - x_{k-1}) \right) dt_{n-1} \\ &= f[x_o, x_1, \dots, x_{n-1}], \end{aligned} \tag{2.105}$$

we will show (2.104) holds for the integer n . Let

$$w = t_n(x_n - x_{n-1}) + \cdots + t_1(x_1 - x_o) + x_o$$

be the new variable. Hence

$$dt_n = \frac{dw}{x_n - x_{n-1}}$$

for $x_n \neq x_{n-1}$. If $t_n = 0$, then $w = w_o$, where

$$w_o = t_{n-1}(x_{n-1} - x_{n-2}) + \cdots + t_1(x_1 - x_o) + x_o.$$

Similarly, if $t_n = t_{n-1}$, then $w = w_1$, where

$$w_1 = t_{n-1}(x_n - x_{n-2}) + \cdots + t_1(x_1 - x_o) + x_o.$$

Now applying the induction hypothesis, we have

$$\begin{aligned}
 & \int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} f^{(n)} \left(x_o + \sum_{k=1}^n t_k (x_k - x_{k-1}) \right) dt_n \\
 &= \int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-2}} \frac{f^{(n-1)}(w_1) - f^{(n-1)}(w_o)}{x_n - x_{n-1}} dt_{n-1} \\
 &= \frac{f[x_o, x_1, \dots, x_{n-2}, x_n] - f[x_o, x_1, \dots, x_{n-2}, x_{n-1}]}{x_n - x_{n-1}} \\
 &= f[x_o, x_1, \dots, x_n]
 \end{aligned}$$

This completes the proof of the theorem.

From the above integral representation, we see that the integrand is a continuous function of the variables x_o, x_1, \dots, x_n , and therefore the left side, $f[x_o, x_1, \dots, x_n]$, is also a continuous function of these variables. If $f(x)$ has a continuous n^{th} derivative, then the above integral representation defines uniquely the continuous extension of $f[x_o, x_1, \dots, x_n]$. For example, if $n = 1$, then the continuous extension of $f[x_o, x_1]$ is

$$f[x_o, x_1] = \begin{cases} \frac{f(x_o) - f(x_1)}{x_o - x_1} & \text{if } x_1 \neq x_o \\ f'(x_o) & \text{if } x_1 = x_o \end{cases}$$

provided $f(x)$ has the first derivative. Because of this unique extension now we can allow some of the nodes, that is x_o, x_1, \dots, x_n , to coalesce if f is suitably differentiable.

Now we present the mean value theorem for divided differences.

Theorem 2.10 *Let $f : [a, b] \rightarrow \mathbb{R}$ be a real valued function with continuous n^{th} derivative and x_o, x_1, \dots, x_n in $[a, b]$. Then there exists a point η in the interval $[\min\{x_o, x_1, \dots, x_n\}, \max\{x_o, x_1, \dots, x_n\}]$ such that*

$$f[x_o, x_1, \dots, x_n] = \frac{f^{(n)}(\eta)}{n!}.$$

Proof: Since $f^{(n)}(x)$ is continuous on $[a, b]$, the function $f^{(n)}(x)$ has a maximum and a minimum on $[a, b]$. Let

$$m = \min f^{(n)}(x) \quad \text{and} \quad M = \max f^{(n)}(x).$$

Then from the integral representation of $f[x_0, x_1, \dots, x_n]$, we have

$$m \prod_{k=1}^n \int_0^{t_{k-1}} dt_k \leq f[x_0, x_1, \dots, x_n] \leq M \prod_{k=1}^n \int_0^{t_{k-1}} dt_k,$$

where $t_0 = 1$. Using the fact that

$$\prod_{k=1}^n \int_0^{t_{k-1}} dt_k = \int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n = \frac{1}{n!},$$

we obtain from the above inequalities

$$m \leq f[x_0, x_1, \dots, x_n] (n!) \leq M.$$

Since $f^{(n)}(x)$ is continuous, by applying the intermediate value theorem to it, we have

$$f[x_0, x_1, \dots, x_n] (n!) = f^{(n)}(\eta)$$

for some $\eta \in [\min\{x_0, x_1, \dots, x_n\}, \max\{x_0, x_1, \dots, x_n\}]$. This yields the asserted result

$$f[x_0, x_1, \dots, x_n] = \frac{f^{(n)}(\eta)}{n!}$$

and the proof of the theorem is now complete.

The mean value theorem for divided differences can be used for defining the functional means. We have seen from our discussion that some of the nodes in the divided differences can coalesce if f is suitably differentiable. For example, if f is differentiable, then

$$f[b, b, a, a] = \frac{f'(b) - 2f[b, a] + f'(a)}{(b-a)^2}.$$

To see this consider,

$$\begin{aligned} f[b, b, a, a] &= \frac{f[b, b, a] - f[b, a, a]}{b-a} \\ &= \frac{1}{b-a} [f[b, b, a] - f[b, a, a]] \\ &= \frac{1}{b-a} \left[\frac{f[b, b] - f[b, a]}{b-a} - \frac{f[b, a] - f[a, a]}{b-a} \right] \\ &= \frac{1}{(b-a)^2} [f[b, b] - 2f[b, a] + f[a, a]] \\ &= \frac{f'(b) - 2f[b, a] + f'(a)}{(b-a)^2}. \end{aligned}$$

For notational simplicity, we denote $f[b, b, a, a]$ by $f[b^{[2]}, a^{[2]}]$. Similarly, in general

$$f[b^{[n]}, a^{[n]}] = f[b, b, \dots, b, a, a, \dots, a].$$

In the above divided difference a and b appear exactly n times each.

Suppose f is $(2n - 1)$ times continuously differentiable in the interval $[a, b]$. Further, we assume that $f^{(2n-1)}(x)$ is strictly monotone in $[a, b]$. Then by the mean value theorem for divided differences, there exists one point $\eta \in [a, b]$ such that

$$f[b^{[n]}, a^{[n]}] = \frac{f^{(2n-1)}(\eta)}{(2n-1)!}.$$

Note that the strict monotonicity of $f^{(2n-1)}(x)$ forces η to be a mean value, that is, $a < \eta < b$. Further, since $f^{(2n-1)}(x)$ is strictly monotone, such a η is also unique, and this defines a *functional mean* $M_f^n(a, b)$ in a and b . A detailed account of this functional mean can be found in Horwitz (1995). Hence

$$M_f^n(a, b) = \left(f^{(2n-1)}\right)^{-1} \left\{ (2n-1)! f[b^{[n]}, a^{[n]}] \right\}.$$

Note that in the above formula $\left(f^{(2n-1)}\right)^{-1}$ denotes the inverse function of $f^{(2n-1)}$. If $n = 1$, then $M_f^n(a, b)$ reduces

$$M_f(a, b) = (f')^{-1} \left(\frac{f(b) - f(a)}{b - a} \right),$$

a functional mean, introduced by Stolarsky (1975) and Mays (1983). Now we give two examples illustrating how well known means can be generated using the functional mean $M_f^n(a, b)$.

Example 2.8 If $f(x) = x^m$, where m is a positive integer greater than or equal to n , then $M_f^n(a, b) = \frac{a+b}{2}$.

If $f(x) = x^m$, then it can be shown that

$$f[x_0, x_1, \dots, x_{m-1}] = x_0 + x_1 + \dots + x_{m-1}. \quad (2.106)$$

Letting $m = 2n$ and

$$x_0 = x_1 = \dots = x_{n-1} = a \quad \text{and} \quad x_n = x_{n+1} = \dots = x_{2n-1} = b$$

in (2.106), we get

$$f[b^{[n]}, a^{[n]}] = n(a + b).$$

Therefore

$$\begin{aligned} M_f^n(a, b) &= (f^{(2n-1)})^{-1} \{(2n-1)! f[b^{[n]}, a^{[n]}]\} \\ &= \frac{1}{(2n)!} (2n-1)! n(a + b) \\ &= \frac{a+b}{2}. \end{aligned}$$

Example 2.9 If $f(x) = \frac{1}{x}$, then $M_f^n(a, b) = \sqrt{ab}$.

If $f(x) = \frac{1}{x}$, then it can be shown that

$$f[x_0, x_1, \dots, x_{n-1}] = \frac{(-1)^n}{x_0 x_1 \cdots x_{n-1}}. \quad (2.107)$$

Hence

$$f[b^{[n]}, a^{[n]}] = \frac{(-1)^{2n-1}}{a^n b^n}.$$

Since $f^{(2n-1)}(x) = \frac{(-1)^{2n-1}(2n-1)!}{x^{2n}}$, thus we have

$$\begin{aligned} M_f^n(a, b) &= (f^{(2n-1)})^{-1} \{(2n-1)! f[b^{[n]}, a^{[n]}]\} \\ &= (a^n b^n)^{\frac{1}{2n}} \\ &= \sqrt{ab}. \end{aligned}$$

Let $f(x) = x^p$, where $p \in \mathbb{R}$. The first example shows that, if p is a positive integer greater than or equal to n , then the functional mean $M_f^n(a, b)$ is the arithmetic mean of a and b . The second example illustrates that if $p = -1$, then the functional mean $M_f^n(a, b)$ is the geometric mean of a and b . Horwitz (1995) has shown that

$$\lim_{n \rightarrow \infty} M_f^n(a, b) = \sqrt{ab}$$

where $f(x) = x^p$ with $p \in \mathbb{R}$. This result says that if f is a power function, then the functional mean asymptotically tends to the geometric mean. It

would be nice to know for what other functions besides the power function the functional mean asymptotically tends to the geometric mean.

As point of our departure, we point out that there are other means which arise as the limiting case of $M_f^n(a, b)$ as $n \rightarrow \infty$. For example, if $f(x) = e^x$, then

$$\lim_{n \rightarrow \infty} M_f^n(a, b) = \frac{a + b}{2}.$$

2.5 Limiting Behavior of Mean Values

If x is a number in the interval $]a, b[$ then applying Lagrange's mean value theorem to the interval $[a, x]$, it is possible to choose a number η_x in $]a, x[$ as a function of x such that

$$f[a, x] = f'(\eta_x). \quad (2.108)$$

In this section, we examine the behavior of the mean value η_x as x approaches the left end point a of the interval $[a, x]$. This type of behavior of integral mean values was first studied by Jacobson (1982) and then by Bao-lin (1997). The behavior of the differential mean values was studied by Powers, Riedel and Sahoo (1998). First, we prove two technical result which will be needed for the main result of this section.

The following theorem gives an alternate representation for the n -point divided difference $f[x_1, x_2, \dots, x_n]$ of f .

Theorem 2.11 *The n -point divided difference of f can be expressed as*

$$f[x_1, x_2, \dots, x_n] = \sum_{j=1}^n \frac{f(x_j)}{\prod_{\substack{k \neq j \\ k=1}}^{n-1} (x_j - x_k)} \quad (2.109)$$

for all positive integers n .

Proof. The proof will be based on induction on n . The expression is trivially true for $n = 1$ and is easy to establish for $n = 2$. Assuming it is true for n , we establish it for $(n + 1)$, using the recursive form of the definition of divided difference as a starting point. We have, by the

definition

$$f[x_1, x_2, \dots, x_{n+1}] = \frac{1}{x_1 - x_{n+1}} (f[x_1, x_2, \dots, x_n] - f[x_2, x_3, \dots, x_{n+1}]).$$

By the induction hypothesis, the right side of the above expression yields

$$\frac{1}{x_1 - x_{n+1}} \left[\sum_{j=1}^n f(x_j) \prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k} - \sum_{j=2}^{n+1} f(x_j) \prod_{\substack{k=2 \\ k \neq j}}^{n+1} \frac{1}{x_j - x_k} \right].$$

Separating the terms based on the left and right endpoints from the above expansion, we get

$$\begin{aligned} & \frac{f(x_1)}{x_1 - x_{n+1}} \prod_{k=2}^n \frac{1}{x_1 - x_k} + \frac{1}{x_1 - x_{n+1}} \sum_{j=2}^n f(x_j) \prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k} \\ & - \frac{f(x_{n+1})}{x_1 - x_{n+1}} \prod_{k=2}^n \frac{1}{x_{n+1} - x_k} - \frac{1}{x_1 - x_{n+1}} \sum_{j=2}^{n+1} f(x_j) \prod_{\substack{k=2 \\ k \neq j}}^{n+1} \frac{1}{x_j - x_k}. \end{aligned}$$

Rearranging, we get

$$\begin{aligned} & \frac{f(x_1)}{x_1 - x_{n+1}} \prod_{k=2}^n \frac{1}{x_1 - x_k} + \frac{f(x_{n+1})}{x_{n+1} - x_1} \prod_{k=2}^n \frac{1}{x_{n+1} - x_k} \\ & + \sum_{j=2}^n \frac{f(x_j)}{x_1 - x_{n+1}} \left[\prod_{\substack{k=1 \\ k \neq j}}^n \frac{1}{x_j - x_k} - \prod_{\substack{k=2 \\ k \neq j}}^{n+1} \frac{1}{x_j - x_k} \right]. \end{aligned}$$

Combining and factoring, we have

$$\begin{aligned} & f(x_1) \prod_{k=2}^{n+1} \frac{1}{x_1 - x_k} + f(x_{n+1}) \prod_{k=1}^n \frac{1}{x_{n+1} - x_k} \\ & + \sum_{j=2}^n \frac{f(x_j)}{x_1 - x_{j+1}} \left[\frac{1}{x_j - x_1} - \frac{1}{x_j - x_{n+1}} \right] \prod_{\substack{k=2 \\ k \neq j}}^n \frac{1}{x_j - x_k}. \end{aligned}$$

The third term can be condensed, yielding

$$\begin{aligned} f(x_1) \prod_{k=2}^{n+1} \frac{1}{x_1 - x_k} + f(x_{n+1}) \prod_{k=1}^n \frac{1}{x_{n+1} - x_k} \\ + \sum_{j=2}^n \frac{f(x_j)}{(x_j - x_1)(x_j - x_{n+1})} \prod_{\substack{k=2 \\ k \neq j}}^n \frac{1}{x_j - x_k}. \end{aligned}$$

This gives

$$f(x_1) \prod_{k=2}^{n+1} \frac{1}{x_1 - x_k} + \sum_{j=2}^{n+1} f(x_j) \prod_{\substack{k=1 \\ k \neq j}}^{n+1} \frac{1}{x_j - x_k} + f(x_{n+1}) \prod_{k=1}^n \frac{1}{x_{n+1} - x_k}$$

and we have established the relationship

$$f[x_1, x_2, \dots, x_{n+1}] = \sum_{j=1}^{n+1} f(x_j) \prod_{\substack{k=1 \\ k \neq j}}^{n+1} \frac{1}{x_j - x_k}.$$

The following result will also be needed for our main result.

Theorem 2.12 *Suppose that $f(x) = x^\ell$ for some nonnegative integer ℓ , then*

$$f[x_1, \dots, x_n] = \begin{cases} 0 & \text{for } \ell < n-1, \\ 1 & \text{for } \ell = n-1, \\ x_1 + \dots + x_n & \text{for } \ell = n \end{cases}$$

for all positive integers n .

Proof: Let $f(x) = x^\ell$, where ℓ is a natural number. We would like to evaluate $f[x_1, x_2, \dots, x_n]$. First consider

$$\begin{aligned} f[x_1, x_2] &= \frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{x_1^\ell - x_2^\ell}{x_1 - x_2} \\ &= \frac{(x_1 - x_2) \sum_{k=0}^{\ell-1} x_1^k x_2^{\ell-1-k}}{x_1 - x_2} \\ &= \sum_{k=0}^{\ell-1} x_1^k x_2^{\ell-1-k} = \sum_{p_1+p_2=\ell-1} x_1^{p_1} x_2^{p_2}, \end{aligned}$$

where p_1 and p_2 are nonnegative integers.

Next, we consider

$$\begin{aligned}
& f[x_1, x_2, x_3] \\
&= \frac{f[x_1, x_3] - f[x_2, x_3]}{x_1 - x_2} \\
&= \frac{\sum_{p_1+p_3=\ell-1} x_1^{p_1} x_3^{p_3} - \sum_{p_2+p_3=\ell-1} x_2^{p_2} x_3^{p_3}}{x_1 - x_2} \\
&= \frac{1}{x_1 - x_2} [(x_1 - x_2)x_3^{\ell-2} + (x_1^2 - x_2^2)x_3^{\ell-3} + \cdots + (x_1^{\ell-2} - x_2^{\ell-2})x_3] \\
&= x_3^{\ell-2} + (x_1 + x_2)x_3^{\ell-3} + (x_1^2 + x_1x_2 + x_2^2)x_3^{\ell-4} \\
&\quad + \cdots + \sum_{p_1+p_2=\ell-2} x_1^{p_1} x_2^{p_2} x_3 \\
&= \sum_{p_2+p_3=\ell-2} x_1^{p_1} x_2^{p_2} x_3^{p_3},
\end{aligned}$$

where p_1, p_2 and p_3 are nonnegative integers. Similarly, it can be shown that

$$f[x_1, x_2, \dots, x_k] = \sum_{p_1+\dots+p_k=\ell-k+1} x_1^{p_1} x_2^{p_2} \cdots x_k^{p_k},$$

where p_1, p_2, \dots, p_k are nonnegative integers. Hence

$$f[x_1, x_2, \dots, x_\ell] = \sum_{p_1+\dots+p_\ell=1} x_1^{p_1} x_2^{p_2} \cdots x_\ell^{p_\ell} = \sum_{j=1}^{\ell-1} x_j,$$

$$f[x_1, x_2, \dots, x_{\ell-1}] = \sum_{p_1+\dots+p_{\ell-1}=0} x_1^{p_1} x_2^{p_2} \cdots x_{\ell-1}^{p_{\ell-1}} = 1,$$

and

$$f[x_1, x_2, \dots, x_{\ell-2}] = 0.$$

This completes the proof of the theorem.

Consider the function $f(t) = t^2$ on the interval $[1, 2]$. Applying Lagrange's mean value theorem to f on the interval $[1, x]$, where $x \in]1, 2[$, we obtain

$$\frac{f(x) - f(1)}{x - 1} = f'(\eta_x)$$

for some η_x in $]1, x[$. Since $f(t) = t^2$, the mean value η_x is given by

$$\eta_x = \frac{1}{2}(x + 1).$$

Now evaluating the limit of $\frac{\eta_x - 1}{x - 1}$ as x approaches 1 from the right, we get

$$\lim_{x \rightarrow 1^+} \frac{\eta_x - 1}{x - 1} = \lim_{x \rightarrow 1^+} \frac{\frac{1}{2}(x + 1) - 1}{x - 1} = \frac{1}{2}.$$

Similarly, if we consider another function $f(t) = e^t$ on the interval $[0, 2]$, then again we have

$$\begin{aligned} \lim_{x \rightarrow 0^+} \frac{\eta_x - 0}{x - 0} &= \lim_{x \rightarrow 0^+} \left[\frac{1}{x} \ln \left(\frac{e^x - 1}{x} \right) \right] \\ &= \lim_{x \rightarrow 0^+} \left[\frac{1}{x} \ln \left(1 + \frac{x}{2} + \frac{x^2}{3!} + \dots \right) \right] = \frac{1}{2}. \end{aligned}$$

These two examples indicate that as x approaches the left end point of the interval from the right, the mean value η_x approaches the midpoint between x and the left end point of the interval. This is true for many functions by the following theorem.

Theorem 2.13 *Suppose the function f is continuously differentiable on $[a, b]$ and twice differentiable at a with $f''(a) \neq 0$. If η_x denotes the mean value in (2.108), then*

$$\lim_{x \rightarrow a^+} \frac{\eta_x - a}{x - a} = \frac{1}{2}.$$

Proof: To establish this theorem, we evaluate

$$\lim_{x \rightarrow a^+} \frac{f(x) - f(a) - (x - a)f'(a)}{(x - a)^2}$$

in two different ways. First, using the mean value theorem, we get

$$\begin{aligned}
 \lim_{x \rightarrow a^+} \frac{f(x) - f(a) - (x - a)f'(a)}{(x - a)^2} &= \lim_{x \rightarrow a^+} \frac{(x - a)f'(\eta_x) - (x - a)f'(a)}{(x - a)^2} \\
 &= \lim_{x \rightarrow a^+} \frac{f'(\eta_x) - f'(a)}{x - a} \\
 &= \lim_{x \rightarrow a^+} \frac{f'(\eta_x) - f'(a)}{\eta_x - a} \lim_{x \rightarrow a^+} \frac{\eta_x - a}{x - a} \\
 &= f''(a) \lim_{x \rightarrow a^+} \frac{\eta_x - a}{x - a}.
 \end{aligned}$$

Then applying L'Hospital rule, we get

$$\lim_{x \rightarrow a^+} \frac{f(x) - f(a) - (x - a)f'(a)}{(x - a)^2} = \lim_{x \rightarrow a^+} \frac{f'(x) - f'(a)}{2(x - a)} = \frac{1}{2} f''(a).$$

Since $f''(a) \neq 0$, we obtain

$$\lim_{x \rightarrow a^+} \frac{\eta_x - a}{x - a} = \frac{1}{2}$$

and the proof is now complete.

The main idea for the above proof came from a paper due to Jacobson (1982). Next, we give a more general theorem which includes the above theorem as a special case. Recall that the mean value theorem for the n -point divided difference states that if $f : [a, b] \rightarrow \mathbb{R}$ is $(n - 1)$ -times continuously differentiable and x_1, \dots, x_n are n distinct points in $[a, b]$, then there exists $\eta \in [\min\{x_1, \dots, x_n\}, \max\{x_1, \dots, x_n\}]$ such that

$$[x_1, \dots, x_n; f(x)] = \frac{f^{(n-1)}(\eta)}{(n-1)!}. \quad (2.110)$$

Here we have used the notation $[x_1, \dots, x_n; f(x)]$ to denote the n -point divided difference, $f[x_1, \dots, x_n]$, of f . We will use this new notation throughout this section only. Further, if $\alpha \in \mathbb{R}$ and $n \in \mathbb{N}$, then the generalized binomial coefficient $\binom{\alpha}{k}$ is defined as

$$\binom{\alpha}{k} = \frac{\alpha(\alpha - 1)\cdots(\alpha - k + 1)}{k!} = \frac{1}{k!} \prod_{i=0}^{k-1} (\alpha - i).$$

In the case where $k = 0$, the product $\prod_{i=0}^{k-1}(\alpha - i)$ is understood to be 1.

We now consider a variable interval $[a, a + x]$ where $0 < x < b - a$. Let $0 \leq m_1 < \dots < m_n \leq 1$. Then $x_1 = a + m_1x, x_2 = a + m_2x, \dots, x_n = a + m_nx$ are n distinct points in $[a, a + x]$. If $f : [a, b] \rightarrow \mathbb{R}$ is $(n-1)$ -times continuously differentiable, then, by the mean value theorem for n -point divided difference applied to f on the interval $[a + m_1x, a + m_nx]$, there exists a mean value in the interval $[a + m_1x, a + m_nx]$ such that (2.110) is satisfied. To emphasize the dependence on the variable x we denote this mean value by η_x . There may be many possible choices for η_x . Therefore, the correspondence $x \rightarrow \eta_x$ involves a choice function. Our goal is to study the behavior of η_x as x goes to zero.

The following theorem is due to Powers, Riedel and Sahoo (1998).

Theorem 2.14 *Let $f : [a, b] \rightarrow \mathbb{R}$ be $(n-1)$ -times continuously differentiable on $[a, b]$ such that*

$$f(t) = p(t) + (t - a)^\alpha g(t), \quad (2.111)$$

where $p(t)$ is a polynomial of degree at most $n-1$, $g^{(n-1)}(t)$ is bounded on $[a, b]$ and $g(a) \neq 0$, and $\alpha \in \mathbb{R} \setminus \{0, 1, \dots, n-1\}$. Then

$$\lim_{x \rightarrow 0^+} \frac{\eta_x - a}{x} = \left[\frac{[m_1, \dots, m_n; x^\alpha]}{\binom{\alpha}{n-1}} \right]^{\frac{1}{\alpha+1-n}} \quad (2.112)$$

where $0 \leq m_1 < \dots < m_n \leq 1$, η_x is the mean value given in (2.110) for $[a + m_1x, \dots, a + m_nx; f(x)]$, and $0 < x < b - a$.

Proof. In view of (2.109), we can write

$$[a + m_1x, \dots, a + m_nx; f(t)] = \sum_{i=1}^n \frac{f(a + m_i x)}{x^{n-1} \prod_{\substack{j \neq i \\ j=1}}^{n-1} (m_i - m_j)}. \quad (2.113)$$

Since $p(t)$ is a polynomial of degree at most $(n-1)$, the left and right side of (2.110) yield the same constant and they cancel, thus we may assume

without loss of generality that $p(t) = 0$. Thus from (2.111), we have

$$f(a + m_i x) = (m_i x)^\alpha g(a + m_i x).$$

Substituting this into the right-hand-side of equation (2.113), we get

$$[a + m_1 x, \dots, a + m_n x; f(t)] = \sum_{i=1}^n \frac{(m_i x)^\alpha g(a + m_i x)}{x^{n-1} \prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)}$$

which is

$$[a + m_1 x, \dots, a + m_n x; f(t)] = x^{\alpha-(n-1)} \sum_{i=1}^n \frac{m_i^\alpha g(a + m_i x)}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)}. \quad (2.114)$$

First, we use the product rule of derivative to determine $f^{(n-1)}(t)$ and then substitute $t = \eta_x$ to obtain

$$f^{(n-1)}(\eta_x) = \sum_{j=0}^{n-1} \binom{n-1}{j} \left\{ \prod_{i=0}^{j-1} (\alpha - i) \right\} (\eta_x - a)^{\alpha-j} g^{(n-1-j)}(\eta_x). \quad (2.115)$$

In the case where $j = 0$, the product $\prod_{i=0}^{j-1} (\alpha - i)$ is understood to be 1. Using equations (2.114) and (2.115) we can rewrite equation (2.110) as follows

$$\begin{aligned} & x^{\alpha-(n-1)} \sum_{i=1}^n \frac{m_i^\alpha g(a + m_i x)}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} \\ &= \frac{1}{(n-1)!} \sum_{j=0}^{n-1} \binom{n-1}{j} \left\{ \prod_{i=0}^{j-1} (\alpha - i) \right\} (\eta_x - a)^{\alpha-j} g^{(n-1-j)}(\eta_x). \end{aligned}$$

Next we divide the above by $x^{\alpha-(n-1)}$ to get

$$\begin{aligned} & \sum_{i=1}^n \frac{m_i^\alpha g(a + m_i x)}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} \\ &= \sum_{j=0}^{n-1} \binom{n-1}{j} \left\{ \frac{\prod_{i=0}^{j-1} (\alpha - i)}{(n-1)!} \right\} \frac{(\eta_x - a)^{\alpha-j}}{x^{\alpha-(n-1)}} g^{(n-1-j)}(\eta_x). \end{aligned} \quad (2.116)$$

Now we can take the limit as x approaches 0 from the right and observe that η_x tends to a . It follows from the hypotheses on g that each summand on the right-hand side of (2.116) is zero except when $j = n-1$. Thus

$$\frac{\prod_{i=0}^{n-2} (\alpha - i)}{(n-1)!} \lim_{x \rightarrow 0^+} \left[g(\eta_x) \left(\frac{\eta_x - a}{x} \right)^{\alpha-(n-1)} \right] = \sum_{i=1}^n \frac{m_i^\alpha \lim_{x \rightarrow 0^+} g(a + m_i x)}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)}.$$

Using the fact that $g(a) \neq 0$ we finally obtain

$$\lim_{x \rightarrow 0^+} \left(\frac{\eta_x - a}{x} \right)^{\alpha-(n-1)} = \frac{1}{\binom{\alpha}{n-1}} \sum_{i=1}^n \frac{m_i^\alpha}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)},$$

where $\binom{\alpha}{n-1}$ represents the generalized binomial coefficient. The desired limit is obtained after taking roots.

To establish the analog of Theorem 2.13 for divided differences we need a version of Taylor's formula which contains the remainder in the Peano form (see Uvarov (1988)).

Theorem 2.15 *Suppose that f is n times differentiable at a , then there*

is a function $\epsilon(x)$ such that

$$f(a+x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} x^k + \epsilon(x) x^n, \quad (2.117)$$

where $\lim_{x \rightarrow 0} \epsilon(x) = 0$.

Using Theorem 2.15 and following the proof of Theorem 2.14, one can establish the following theorem. However, we present the entire proof for the sake of completeness.

Theorem 2.16 Suppose $f : [a, b] \rightarrow \mathbb{R}$ has a continuous $(n-1)^{\text{st}}$ derivative and is $k \geq n$ times differentiable at a with $f^{(i)}(a) = 0$ for $i = n, \dots, (k-1)$ (obviously if $k = n$ this condition is vacuous), and $f^{(k)}(a) \neq 0$. Then

$$\lim_{x \rightarrow 0^+} \frac{\eta_x - a}{x} = \left[\frac{[m_1, \dots, m_n; x^k]}{\binom{k}{n-1}} \right]^{\frac{1}{k+1-n}} \quad (2.118)$$

where $0 \leq m_1 < \dots < m_n \leq 1$, η_x is the mean value given in (2.110) for $[a + m_1 x, \dots, a + m_n x; f(x)]$, and $0 < x < b - a$.

Proof. In view of (2.109), we can write

$$[a + m_1 x, \dots, a + m_n x; f(t)] = \sum_{i=1}^n \frac{f(a + m_i x)}{x^{n-1} \prod_{\substack{j \neq i \\ j=1}}^{n-1} (m_i - m_j)}, \quad (2.119)$$

and using Theorem 2.15, we expand $f(a + m_i x)$ as

$$f(a + m_i x) = \sum_{\ell=0}^k \frac{f^{(\ell)}(a)}{\ell!} (m_i x)^\ell + \epsilon(m_i x) (m_i x)^k,$$

where $\lim_{x \rightarrow 0^+} \epsilon(m_i x) = 0$. Next, using the hypothesis that $f^{(i)}(a) = 0$ for $i = n, \dots, k-1$ in the above, we obtain

$$f(a + m_i x) = \sum_{\ell=0}^{n-1} \frac{f^{(\ell)}(a)}{\ell!} (m_i x)^\ell + \frac{f^{(k)}(a)}{k!} (m_i x)^k + \epsilon(m_i x) (m_i x)^k.$$

Substituting this into the right-hand-side of equation (2.119) yields

$$[a + m_1x, \dots, a + m_nx; f(t)]$$

$$= \sum_{i=1}^n \left\{ \frac{\sum_{\ell=0}^{n-1} \frac{f^{(\ell)}(a)}{\ell!} (m_i x)^\ell + \frac{f^{(k)}(a)}{k!} (m_i x)^k + \epsilon(m_i x) (m_i x)^k}{x^{n-1} \prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} \right\}.$$

Now we can rearrange these terms to obtain expressions corresponding to divided differences in the m_i :

$$[a + m_1x, \dots, a + m_nx; f(t)]$$

$$\begin{aligned} &= \sum_{\ell=0}^{n-2} \frac{f^{(\ell)}(a)}{\ell! x^{(n-1)-\ell}} \sum_{i=1}^n \frac{m_i^\ell}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} + \frac{f^{(n-1)}(a)}{(n-1)!} \sum_{i=1}^n \frac{m_i^{n-1}}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} \\ &\quad + \frac{f^{(k)}(a)}{k!} \sum_{i=1}^n \frac{m_i^k x^{k-(n-1)}}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} + \sum_{i=1}^n \frac{\epsilon(m_i x) x^k m_i^k}{x^{n-1} \prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)}. \end{aligned}$$

With the help of Theorem 2.12, this yields

$$[a + m_1x, \dots, a + m_nx; f(t)]$$

$$\begin{aligned} &= \frac{f^{(n-1)}(a)}{(n-1)!} + \frac{f^{(k)}(a)}{k!} \sum_{i=1}^n \frac{m_i^k x^{k-(n-1)}}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} \\ &\quad + \sum_{i=1}^n \frac{\epsilon(m_i x) x^k m_i^k}{x^{n-1} \prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)}. \end{aligned} \tag{2.120}$$

Expanding $f^{(n-1)}(x)$ into a Taylor polynomial of degree $k - (n - 1)$ with remainder and evaluating at the mean value η_x , we have

$$f^{(n-1)}(\eta_x) = \sum_{\ell=0}^{k-(n-1)} \frac{f^{(\ell+n-1)}(a)}{\ell!} (\eta_x - a)^\ell + \hat{\epsilon}(\eta_x - a) (\eta_x - a)^{k-(n-1)},$$

where $\lim_{\eta_x \rightarrow a} \hat{\epsilon}(\eta_x - a) = 0$. Using the hypotheses of the theorem and the above equation, the right hand side of equation (2.110) becomes

$$\begin{aligned} \frac{f^{(n-1)}(\eta_x)}{(n-1)!} &= \frac{f^{(n-1)}(a)}{(n-1)!} + \frac{f^{(k)}(a)}{(n-1)!(k-(n-1))!} (\eta_x - a)^{k-(n-1)} \\ &\quad + \frac{\hat{\epsilon}(\eta_x - a) (\eta_x - a)^{k-(n-1)}}{(n-1)!}. \end{aligned} \tag{2.121}$$

Using (2.110), (2.120) and (2.121), and canceling $\frac{f^{(n-1)}(a)}{(n-1)!}$, we get

$$\begin{aligned} &\frac{f^{(k)}(a)}{k!} \sum_{i=1}^n \frac{m_i^k x^{k-(n-1)}}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} + \sum_{i=1}^n \frac{\epsilon(m_i x) x^k m_i^k}{x^{n-1} \prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} \\ &= \frac{f^{(k)}(a)}{(n-1)!(k-(n-1))!} (\eta_x - a)^{k-(n-1)} + \frac{\hat{\epsilon}(\eta_x - a) (\eta_x - a)^{k-(n-1)}}{(n-1)!}. \end{aligned}$$

Using $x > 0$ and the fact that $f^{(k)}(a) \neq 0$, we obtain

$$\begin{aligned} \left(\frac{\eta_x - a}{x} \right)^{k-(n-1)} &= \frac{(n-1)!(k-(n-1))!}{k!} \sum_{i=1}^n \frac{m_i^k}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} + \\ \frac{(n-1)!(k-(n-1))!}{f^{(k)}(a) x^{k-(n-1)}} &\left(\left(\sum_{i=1}^n \frac{\epsilon(m_i x) x^k m_i^k}{x^{n-1} \prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)} - \frac{\hat{\epsilon}(\eta_x - a) (\eta_x - a)^{k-(n-1)}}{(n-1)!} \right) \right). \end{aligned}$$

Since $\lim_{x \rightarrow 0^+} \epsilon(m_i x) = 0$ ($i = 1, \dots, n$) and $\lim_{x \rightarrow 0^+} \tilde{\epsilon}(\eta_x - a) = 0$, it follows that

$$\lim_{x \rightarrow 0^+} \left(\frac{\eta_x - a}{x} \right)^{k-(n-1)} = \frac{(n-1)!(k-(n-1))!}{k!} \sum_{i=1}^n \frac{m_i^k}{\prod_{\substack{j \neq i \\ j=1}}^n (m_i - m_j)},$$

which after taking the roots yields the desired limit. The proof of the theorem is complete.

Suppose f is continuous on the interval $[a, b]$ and is differentiable at a with $f'(a) \neq 0$. For each $x \in]a, b[$, let ξ_x be the value determined by the integral mean value theorem where

$$\int_a^x f(t) dt = f(\xi_x)(x - a).$$

If $F(s) = \int_a^s f(t) dt$, then the previous equation can be written as

$$F'(\xi_x) = \frac{F(x) - F(a)}{x - a}.$$

The last equation can be written in terms of the 2-point divided difference as

$$[a, a+x; F(s)] = F'(\xi_x).$$

Since $F(s)$ is continuously differentiable on $[a, b]$ and is twice differentiable at a with $F''(a) \neq 0$ it follows from Theorem 2.16 (with $m_1 = 0$ and $n = 2$) that

$$\lim_{x \rightarrow 0^+} \frac{\xi_x - a}{x} = \frac{[0, 1; x^2]}{\binom{2}{1}} = \frac{1}{2}.$$

This was the result established by Jacobson (1982) which follows immediately from Theorem 2.16. Similarly, the results obtained by Bao-lin (1997) also follow immediately from Theorem 2.16. Finally, we note that the right-hand side of (2.118) with $n = 2$ is the Stolarsky mean of order k (see Stolarsky (1975)), that is

$$M_f(m_1, m_2) = \left[\left(\frac{1}{k} \right) \frac{m_1^k - m_2^k}{m_1 - m_2} \right]^{\frac{1}{k-1}},$$

where $f(x) = x^k$ (see Example 2.7 in this chapter).

2.6 Cauchy's MVT and Functional Equations

Augustine-Louis Cauchy (1789-1857) gave the following generalization of the mean value theorem which now bears his name.

Theorem 2.17 *For all real-valued functions f and g differentiable on a real interval I and for all pairs $x_1 \neq x_2$ in I , there exists a point η depending on x_1 and x_2 such that*

$$[f(x_1) - f(x_2)] g'(\eta) = [g(x_1) - g(x_2)] f'(\eta). \quad (2.122)$$

Proof: Define $h(x) = [f(x_1) - f(x_2)] g(x) - [g(x_1) - g(x_2)] f(x)$ for all $x \in I$. Then h is differentiable on I and, furthermore, we have

$$h(x_1) = f(x_2)g(x_1) - g(x_2)f(x_1) = h(x_2).$$

By Rolle's theorem there is a $\eta \in]x_1, x_2[$, such that

$$0 = h'(\eta) = [f(x_1) - f(x_2)] g'(\eta) - [g(x_1) - g(x_2)] f'(\eta),$$

which is (2.103) and the proof is now complete.

Cauchy's mean value theorem is used in proving a popular method for calculating the limits of certain ratios of functions. This method is attributed to Guillaume Francois Marquis de l'Hospital (1661-1704) and is known as l'Hospital's rule. In 1696 the Marquis de l'Hospital compiled the lecture notes of his teacher Johann Bernoulli (1667-1748) and it was there the so-called l'Hospital rule first appeared. It is perhaps more accurate to refer to this rule as Bernoulli-l'Hospital's rule. Note that the name l'Hospital follows the old French spelling and the letter s is not to be pronounced. In modern French this name is spelled as l'Hopital.

Like Lagrange's mean value theorem, Cauchy's mean value theorem can also be used for deriving Stolarsky's mean. Let α, β be two distinct nonzero real numbers. Cauchy's mean value theorem applied to function $f(t) = t^\alpha$, $g(t) = t^\beta$ on the interval $[x, y]$ of the positive real numbers yields

$$\eta_{\alpha, \beta}(x, y) = \left(\frac{\beta(x^\alpha - y^\alpha)}{\alpha(x^\beta - y^\beta)} \right)^{\frac{1}{\alpha-\beta}}.$$

This mean value $\eta_{\alpha,\beta}(x, y)$ can be extended continuously to the domain $\{(\alpha, \beta, x, y) \mid \alpha, \beta \in \mathbb{R}, x, y \in]0, \infty[\}$ (see Losonczi and Pales (1998)). This extension is given by

$$\eta_{\alpha,\beta}(x, y) = \begin{cases} \left(\frac{\beta(x^\alpha - y^\alpha)}{\alpha(x^\beta - y^\beta)} \right)^{\frac{1}{\alpha-\beta}} & \text{if } \alpha\beta(\alpha - \beta)(x - y) \neq 0, \\ \left(\frac{(x^\alpha - y^\alpha)}{\alpha(\ln x - \ln y)} \right)^{\frac{1}{\alpha}} & \text{if } \alpha(x - y) \neq 0, \beta = 0, \\ \left(\frac{\beta(\ln x - \ln y)}{x^\beta - y^\beta} \right)^{\frac{1}{\alpha-\beta}} & \text{if } \beta(x - y) \neq 0, \alpha = 0, \\ \exp \left(-\frac{1}{\alpha} + \frac{x^\alpha \ln x - y^\alpha \ln y}{x^\alpha - y^\alpha} \right) & \text{if } \beta(x - y) \neq 0, \alpha = \beta, \\ \sqrt{xy} & \text{if } (x - y) \neq 0, \alpha = \beta = 0, \\ x & \text{if } (x - y) \neq 0. \end{cases}$$

If one asks for what f and g the mean value η depends on x_1 and x_2 in a given manner, then equation (2.122) appears as a functional equation. This functional equation was investigated by Aumann (1936) by assuming

$$\eta(x_1, x_2) = F^{-1} \left(\frac{F(x_1) + F(x_2)}{2} \right), \quad (2.123)$$

where F is a continuous and strictly monotonic function. The function η in equation (2.123) is called a *quasiarithmetic mean* of x_1 and x_2 . If we take $\eta(x, y) = x + y$ and replace the derivatives of f and g by unknown functions h and k respectively, then from equation (2.122) we obtain the following functional equation

$$[f(x) - f(y)] k(x + y) = [g(x) - g(y)] h(x + y) \quad (2.124)$$

for all $x, y \in I$ with $x \neq y$. The condition $x \neq y$ can be dropped from the above equation since the functional equation holds even without this condition.

In section 4, we examined Lagrange's mean value theorem for divided differences. Rätz and Russell (1987) have established Cauchy's mean value theorem for divided differences. Next, we state and prove their result.

Theorem 2.18 *Let $f, g : [a, b] \rightarrow \mathbb{R}$ be real valued functions with continuous n^{th} derivatives and $g^{(n)}(t) \neq 0$ on $[a, b]$. Further, let x_0, x_1, \dots, x_n in $[a, b]$. Then there exists a point $\eta \in [\min\{x_0, x_1, \dots, x_n\}, \max\{x_0, x_1, \dots, x_n\}]$*

such that

$$f[x_o, x_1, \dots, x_n] g^{(n)}(\eta) = g[x_o, x_1, \dots, x_n] f^{(n)}(\eta). \quad (2.125)$$

Proof. Without loss of generality we may assume $x_o \leq x_1 \leq \dots \leq x_n$. If $x_o = x_1 = \dots = x_n$, then from the definition of divided differences and the fact that f and g are n times continuously differentiable, (2.125) holds with $x_o = x_1 = \dots = x_n = \eta$.

Next suppose $x_o < x_n$. For $x_o \leq t \leq x_n$, define

$$F(t) = f[t, x_1, \dots, x_{n-1}] \quad \text{and} \quad G(t) = g[t, x_1, \dots, x_{n-1}]. \quad (2.126)$$

From the definition of divided differences and (2.126), we see that

$$f[x_o, x_1, \dots, x_n] = \frac{F(x_o) - F(x_n)}{x_o - x_n} \quad (2.127)$$

and

$$g[x_o, x_1, \dots, x_n] = \frac{G(x_o) - G(x_n)}{x_o - x_n}. \quad (2.128)$$

Since $g^{(n)}(t) \neq 0$ on $[a, b]$, one can conclude that $g[x_o, x_1, \dots, x_n] \neq 0$. Next, we define

$$H(t) = g[x_o, x_1, \dots, x_n]F(t) - f[x_o, x_1, \dots, x_n]G(t). \quad (2.129)$$

Using (2.127) and (2.128) in (2.129) it is easy to see that

$$H(x_o) = H(x_n). \quad (2.130)$$

The linearity of the devided difference and (2.129) implies that

$$\begin{aligned} H(t) &= g[x_o, x_1, \dots, x_n]F(t) - f[x_o, x_1, \dots, x_n]G(t) \\ &= \left(\frac{G(x_o) - G(x_n)}{x_o - x_n} \right) F(t) - \left(\frac{F(x_o) - F(x_n)}{x_o - x_n} \right) G(t) \\ &= \left(\frac{G(x_o) - G(x_n)}{x_o - x_n} \right) f[t, x_1, \dots, x_{n-1}] \\ &\quad - \left(\frac{F(x_o) - F(x_n)}{x_o - x_n} \right) g[t, x_1, \dots, x_{n-1}] \\ &= h[t, x_1, \dots, x_{n-1}], \end{aligned}$$

where

$$h(t) = g[x_0, x_1, \dots, x_n]f(t) - f[x_0, x_1, \dots, x_n]g(t) \quad (2.131)$$

with $x_0 \leq t \leq x_n$. Differentiating $H(t)$ with respect to t , we have from the properties of divided differences (see Schumaker (1981))

$$H'(t) = h[t, t, x_1, \dots, x_{n-1}]. \quad (2.132)$$

Since f and g are n -times differentiable, so also h . Thus, using the mean value theorem for divided differences, we have

$$h[t, t, x_1, \dots, x_{n-1}] = \frac{h^{(n)}(\xi(t))}{n!} \quad (2.133)$$

for some $\xi(t)$ in the interval $[x_0, x_n]$. Thus, from (2.132) and (2.133), we have

$$H'(t) = \frac{h^{(n)}(\xi(t))}{n!}. \quad (2.134)$$

Since H is differentiable and H satisfies (2.130), we obtain

$$H'(\theta) = 0 \quad (2.135)$$

for some θ in $]x_0, x_n[$. From (2.134) and (2.136) and calling $\xi(\theta)$ to be η , we have

$$\frac{h^{(n)}(\eta)}{n!} = 0. \quad (2.136)$$

Now using (2.131) in (2.136), we see that

$$g[x_0, x_1, \dots, x_n]f^{(n)}(\eta) - f[x_0, x_1, \dots, x_n]g^{(n)}(\eta) = 0. \quad (2.137)$$

Note that $\eta \in [x_0, x_n]$. Now the proof of the theorem is complete.

2.7 Some Open Problems

We conclude this chapter with three open problems. The first open problem is the following. Find the general solution $f, g, h, k : \mathbb{R} \rightarrow \mathbb{R}$ of the functional equation

$$[f(x) - f(y)] k(x + y) = [g(x) - g(y)] h(x + y)$$

for all $x, y \in \mathbb{R}$ with $x \neq y$.

The second open problem is the following. Find the general solution $f, h : \mathbb{R} \rightarrow \mathbb{R}$ of the functional equation

$$f[x_1, x_2, \dots, x_n] = h\left(g^{-1}\left(\frac{1}{n} \sum_{i=1}^n g(x_i)\right)\right),$$

where g is a continuous and strictly monotone function.

The third problem is the following. Let w_1, w_2, \dots, w_n be *a priori* chosen real parameters. Find the general solution of the functional equation

$$f[x_1, x_2, \dots, x_n] = h(w_1 x_1 + w_2 x_2 + \dots + w_n x_n), \quad (2.138)$$

for all $x_i \in \mathbb{R}$ with $x_i \neq x_j$ for $i, j = 1, 2, \dots, n$ and $i \neq j$.

Finally, our last problem is the following. Given two positive numbers a and b , characterize all functions f for which the functional mean $M_f^n(a, b)$ tends to the geometric mean when $n \rightarrow \infty$, that is

$$\lim_{n \rightarrow \infty} M_f^n(a, b) = \sqrt{ab}.$$

In connection with the asymptotic behavior of the functional mean, Horwitz (1995) asked what other means arise when f is not a power function. We have seen that the geometric and arithmetic means do arise as the asymptotic behavior of the functional mean $M_f^n(a, b)$.

Chapter 3

Pompeiu's Mean Value Theorem and Associated Functional Equations

In this chapter, we examine a mean value theorem due to Pompeiu (1946). We also examine a generalization of Pompeiu's mean value theorem proposed by Boggio (1947-48). In Chapter 2, we encountered many functional equations motivated by Lagrange's mean value theorem. Similar functional equations arise from Pompeiu's mean value theorem. These functional equations are known as Stamate type functional equations. Section two deals with some Stamate type functional equations and their generalizations. Section three includes the generalization of Pompeiu's mean value theorem proposed by Boggio (1947-48). In this section, we present a functional equation studied by Kuczma (1991). This section is a bit technical and to some extent incomplete. The interested reader should refer to Kuczma's original paper. In section four, we examine a generalization of the Stamate type equation and then solve some other related functional equations that arise from Simpson's rule for numerically evaluating definite integrals. Finally, we close this chapter with a brief discussion of some problems that need an investigation.

3.1 Pompeiu's Mean Value Theorem

In 1946, Pompeiu derived a variant of Lagrange's mean value theorem, now known as Pompeiu's mean value theorem.

Theorem 3.1 *For every real valued function f differentiable on an interval $[a, b]$ not containing 0 and for all pairs $x_1 \neq x_2$ in $[a, b]$, there exists*

a point ξ in $]x_1, x_2[$ such that

$$\frac{x_1 f(x_2) - x_2 f(x_1)}{x_1 - x_2} = f(\xi) - \xi f'(\xi). \quad (3.1)$$

Proof: Define a real valued function F on the interval $[\frac{1}{b}, \frac{1}{a}]$ by

$$F(x) = tf\left(\frac{1}{t}\right). \quad (3.2)$$

Since f is differentiable on $[a, b]$ and 0 is not in $[a, b]$, we see that F is differentiable on $[\frac{1}{b}, \frac{1}{a}]$ and

$$F'(t) = f\left(\frac{1}{t}\right) - \frac{1}{t}f'\left(\frac{1}{t}\right). \quad (3.3)$$

Applying the mean value theorem to F on the interval $[x, y] \subset [\frac{1}{b}, \frac{1}{a}]$, we get

$$\frac{F(x) - F(y)}{x - y} = F'(\eta) \quad (3.4)$$

for some $\eta \in]x, y[$. Let $x_2 = \frac{1}{x}$, $x_1 = \frac{1}{y}$, and $\xi = \frac{1}{\eta}$. Then since $\eta \in]x, y[$, we have

$$x_1 < \xi < x_2.$$

Now, using (3.2) and (3.3) on (3.4), we have

$$\frac{xf\left(\frac{1}{x}\right) - yf\left(\frac{1}{y}\right)}{x - y} = f\left(\frac{1}{\eta}\right) - \frac{1}{\eta}f'\left(\frac{1}{\eta}\right)$$

that is

$$\frac{x_1 f(x_2) - x_2 f(x_1)}{x_1 - x_2} = f(\xi) - \xi f'(\xi).$$

This completes the proof of the theorem.

Let us discuss the geometrical interpretation of this theorem. The equation of the secant line joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$ is given by

$$y = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1} (x - x_1).$$

This line intersects the y -axis at the point $(0, y)$, where y is

$$\begin{aligned} y &= f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1} (0 - x_1) \\ &= \frac{x_2 f(x_1) - x_1 f(x_2) - x_1 f(x_2) + x_1 f(x_1)}{x_2 - x_1} \\ &= \frac{x_1 f(x_2) - x_2 f(x_1)}{x_1 - x_2}. \end{aligned}$$

The equation of the tangent line at the point $(\xi, f(\xi))$ is

$$y = (x - \xi) f'(\xi) + f(\xi).$$

This tangent line intersects the y -axis at the point $(0, y)$, where

$$y = -\xi f'(\xi) + f(\xi).$$

If this tangent line intersects the y -axis at the same point as the secant line joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$, then we have

$$\frac{x_1 f(x_2) - x_2 f(x_1)}{x_1 - x_2} = f(\xi) - \xi f'(\xi),$$

which is the equation (3.1) in Theorem 3.1. Hence the geometric meaning of this is that the tangent at the point $(\xi, f(\xi))$ intersects on the y -axis at the same point as the secant line connecting the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$. This is illustrated in the Figure 3.1.

3.2 Stamate Type Equations

The algebraic expression (3.1) yields a functional equation. It turns out (see Aczél (1985) and Aczél and Kuczma (1989)) that the exact form of the right-hand side is not essential. The relevant fact is that the right-hand side of (3.1) depends only on ξ and not directly on x_1 and x_2 . Thus, we have the following functional equation

$$\frac{xf(y) - yf(x)}{x - y} = h(\xi(x, y)), \quad \forall x, y \in \mathbb{R} \quad \text{with } x \neq y. \quad (3.5)$$

Similar to divided difference, a variant of it was defined in Chung and Sahoo (1993) recursively as

$$f\{x_1\} = f(x_1),$$

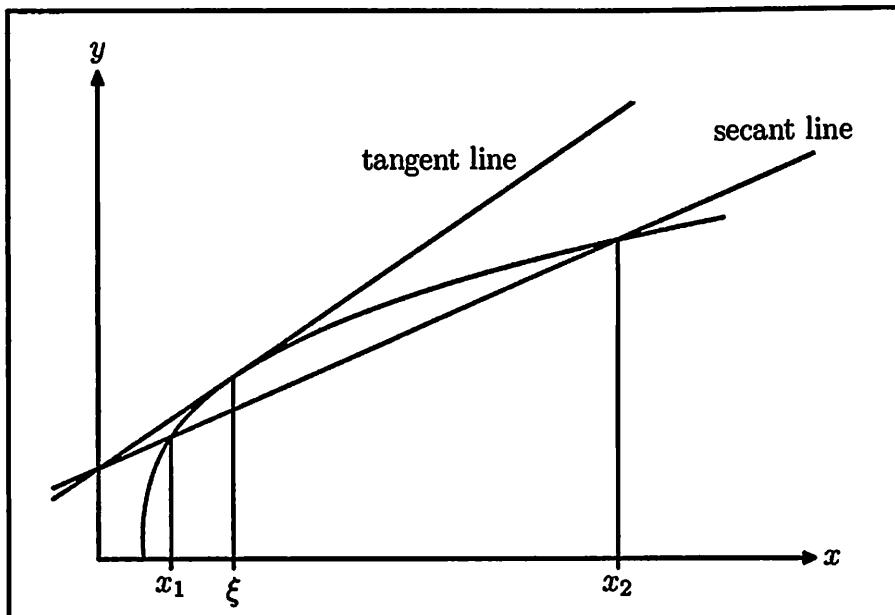


Fig. 3.1 An Interpretation of Pompeiu's Mean Value Theorem.

and

$$f\{x_1, x_2, \dots, x_n\} = \frac{x_n f\{x_1, x_2, \dots, x_{n-1}\} - x_1 f\{x_2, x_3, \dots, x_n\}}{x_1 - x_n}.$$

An easy computation shows that

$$f\{x_1, x_2\} = \frac{x_2 f(x_1) - x_1 f(x_2)}{x_2 - x_1}$$

and

$$f\{x_1, x_2, \dots, x_n\} = \sum_{i=1}^n \left(\prod_{j \neq i}^n \frac{x_j}{x_i - x_j} \right) f(x_i).$$

The following results were established in Aczél and Kuczma (1989).

Theorem 3.2 *The functions $f, h : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$f\{x, y\} = h(x + y) \quad \text{for all } x, y \in \mathbb{R} \text{ with } x \neq y, \quad (3.6)$$

if and only if

$$f(x) = ax + b \quad \text{and} \quad h(x) = b, \quad (3.7)$$

where a, b are arbitrary constants.

Proof: We write (3.6) as

$$x f(y) - y f(x) = (x - y) h(x + y) \quad (3.8)$$

which is now true for all $x, y \in \mathbb{R}$, also for $x = y$. Substituting $y = 0$ in (3.8), we get $xf(0) = xh(x)$, that is

$$h(x) = f(0) = b \quad 0 \neq x \in \mathbb{R}. \quad (3.9)$$

Letting (3.9) into (3.8), we have

$$x f(y) - y f(x) = (x - y) b \quad (3.10)$$

for all $x, y \in \mathbb{R}$ with $x+y \neq 0$. Putting $x = 1$ and $y \neq -1$ (so that $x+y \neq 0$) in (3.10) we obtain

$$f(y) = [f(1) - b]y + b = ay + b \quad (3.11)$$

for all $y \neq -1$. Letting $y = 2$ in (3.11), we see that

$$f(2) = 2f(1) - b. \quad (3.12)$$

Next, putting $x = -1$ and $y = 2$ in (3.8) and then using (3.9) and (3.12), we get

$$f(-1) = -[f(1) - b] + b,$$

that is

$$f(-1) = -a + b. \quad (3.13)$$

Together with (3.13), we see that (3.11) holds for all $y \in \mathbb{R}$. Next, substituting $x = 1$ and $y = -1$ in (3.8), we obtain

$$h(0) = b$$

so that (3.9) holds for all $x \in \mathbb{R}$. Hence we have the asserted solution (3.7) and the proof is now complete.

The following lemma sets the path for a generalization of Theorem 3.2.

Lemma 3.1 *If $f, g, h : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$\frac{xf(y) - yg(x)}{x - y} = h(x + y)$$

for all $x, y \in \mathbb{R}$ with $x \neq y$, then

$$f(x) = g(x), \quad x \in \mathbb{R}.$$

Proof: Interchanging x with y in the above functional equation, we have

$$\frac{yf(x) - xg(y)}{y - x} = h(y + x).$$

Next, comparing with this resulting equation with the functional equation in the lemma, we obtain

$$xf(y) - yg(x) = xg(y) - yf(x), \quad x, y \in \mathbb{R}, \quad x \neq y,$$

whence

$$\frac{f(x) - g(x)}{x} = \frac{f(y) - g(y)}{y}, \quad x, y \in \mathbb{R} \setminus \{0\}, \quad x \neq y.$$

Let α be a fixed nonzero real number and put

$$c = \frac{g(\alpha) - f(\alpha)}{\alpha}.$$

Thus, from the above equation, we have

$$f(x) = g(x) + cx, \quad x \in \mathbb{R} \setminus \{0, \alpha\}.$$

Let $u, v \in \mathbb{R} \setminus \{0, \alpha\}$ with $u \neq v$. Putting $x = u$ and $y = v$ in the above equation, we see that $c = -c$ and hence $c = 0$. Thus

$$f(x) = g(x), \quad x \in \mathbb{R} \setminus \{0\}.$$

Next, letting $x = \alpha$ and $y = 0$ in the functional equation of the lemma, we get $f(0) = h(\alpha)$. Further, putting $x = 0$ and $y = \alpha$ in the functional equation of the lemma, we have $h(\alpha) = g(0)$. Thus, we have $f(0) = g(0)$, and $f(x) = g(x)$ for all $x \in \mathbb{R}$. The proof of the lemma is now complete.

The following corollary follows from Lemma 3.1 and Theorem 3.2.

Corollary 3.1 *The functions $f, h : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$\frac{xf(y) - yg(x)}{x - y} = h(x + y) \quad \text{for all } x, y \in \mathbb{R} \text{ with } x \neq y,$$

if and only if

$$f(x) = g(x) = ax + b \quad \text{and} \quad h(x) = b,$$

where a, b are arbitrary constants.

Next, we present a result similar to Theorem 2.5 in Chapter 2.

Theorem 3.3 *Let s and t be the real parameters. The functions $f, h : \mathbb{R} \rightarrow \mathbb{R}$ satisfy*

$$\frac{xf(y) - yf(x)}{x - y} = h(sx + ty) \quad (3.14)$$

for all $x, y \in \mathbb{R}$, $x \neq y$ if and only if

$$f(x) = ax + b \quad (3.15)$$

$$h(x) = \begin{cases} \text{arbitrary with } b = h(0) & \text{if } s = 0 = t \\ b & \text{if } s = -t, x \neq 0 \\ b & \text{otherwise,} \end{cases} \quad (3.16)$$

where a, b are arbitrary constants.

Proof: Rewriting (3.14), we have

$$x f(y) - y f(x) = (x - y) h(sx + ty) \quad (3.17)$$

for $x, y \in \mathbb{R}$ with $x \neq y$. To establish the theorem, we consider several cases.

Case 1. Suppose $s = 0 = t$. Then from (3.17), we have

$$x [f(y) - b] = y [f(x) - b]. \quad (3.18)$$

Letting $y = 1$ in the above equation, we get

$$f(x) = [f(1) - b]x + b = ax + b, \quad (3.19)$$

where $a = f(1) - b$. Thus, we have the asserted solution

$$\begin{aligned} f(x) &= ax + b \\ h(x) &= \text{arbitrary with } h(0) = b. \end{aligned}$$

Case 2. Suppose $t = 0$ but $s \neq 0$. Then (3.17) yields

$$x f(y) - y f(x) = (x - y) h(sx). \quad (3.20)$$

Letting $y = 0$ in (3.20), we have $x f(0) = x h(sx)$, that is

$$h(x) = b, \quad x \in \mathbb{R} \setminus \{0\}, \quad (3.21)$$

where $b = f(0)$. Using (3.21) in (3.20), we have

$$x [g(y) - b] = y [f(x) - b], \quad x \neq 0. \quad (3.22)$$

Letting $x = 1$ in (3.22), we have

$$f(y) = [f(1) - b] x + b = ay + b \quad y \in \mathbb{R}. \quad (3.23)$$

Letting $x = 0$ in (3.20), we get $h(0) = f(0) = b$ and hence (3.21) holds for all $x \in \mathbb{R}$.

Case 3. Suppose $t \neq 0$ but $s = 0$. This case can be handled similar to the previous case. Thus, we have the solution (3.15)–(3.16) as asserted in the theorem.

Case 4. Suppose $s \neq 0 \neq t$. Letting $y = 0$, we get $x f(0) = x h(sx)$. Hence,

$$h(x) = b, \quad x \in \mathbb{R} \setminus \{0\}, \quad (3.24)$$

where $b = f(0)$. Letting (3.24) into (3.23), we get

$$x f(y) - y f(x) = (x - y) b \quad (3.25)$$

with $sx + yt \neq 0$. Hence putting $x = 1$ in (3.25), we get

$$f(y) = y [f(1) - b] + b = ay + b \quad (3.26)$$

for $y \neq -\frac{s}{t}$. Letting $x = -\frac{s}{t}$ and $y = 2\frac{s}{t}$ in (3.17), we have

$$-\frac{s}{t} f\left(2\frac{s}{t}\right) - 2\frac{s}{t} f\left(-\frac{s}{t}\right) = -3\frac{s}{t} h\left(\frac{s}{t}\right)$$

that is

$$f\left(2\frac{s}{t}\right) + 2 f\left(-\frac{s}{t}\right) = 3b.$$

Therefore, by (3.26), we have

$$f\left(-\frac{s}{t}\right) = -\frac{s}{t}a + b.$$

Thus (3.26) holds for all $y \in \mathbb{R}$.

Next, we show that $h(x) = b$ holds for all $x \in \mathbb{R}$ except the case when $s = -t$. If $s = -t$, then $h(x)$ can be defined arbitrarily at $x = 0$ and $h(x) = b$ for all $x \in \mathbb{R} \setminus \{0\}$.

If $s \neq -t$, then we show that $h(0) = b$. Letting $x = 1$ and $y = -\frac{s}{t}$ in (3.26), we obtain

$$f\left(-\frac{s}{t}\right) + \frac{s}{t}f(1) = \left(1 + \frac{s}{t}\right)h(0),$$

that is

$$h(0) = b$$

and hence we have $h(x) = b$ for all $x \in \mathbb{R}$. This completes the proof of the theorem.

We close this section by pointing out the connection between the two functional equations:

$$\frac{F(x) - F(y)}{x - y} = F'(\xi(x, y))$$

and

$$\frac{xf(y) - yf(x)}{x - y} = f(\xi) - \xi f'(\xi(x, y)).$$

It was remarked in Kuczma (1989) that if one defines

$$\begin{aligned} F(x) &= xf\left(\frac{1}{x}\right) \\ \xi(x, y) &= \frac{1}{\xi\left(\frac{1}{x}, \frac{1}{y}\right)} \end{aligned}$$

for $x \in \mathbb{R} \setminus \{0\}$, then one obtains

$$\frac{F(x) - F(y)}{x - y} = F'(\xi(x, y)).$$

Thus, to solve the first functional equation if 0 is not in the domain, one may apply the above transformation and then solve the second functional equation to obtain the solution of the first functional equation. If 0 is in the domain, then we can solve the second equation on an interval not containing 0 and then check whether the solutions obtained admit extensions to the whole domain satisfying the first equation.

3.3 An Equation of Kuczma

Boggio (1947-48) gave the following generalization of Pompeiu's mean value theorem. We state his theorem without a proof. The interested reader should refer to Boggio (1947-48).

Theorem 3.4 *For all real valued functions f and g differentiable on an interval $[a, b]$ not containing 0 and for all pairs $x_1 \neq x_2$ in $[a, b]$, there exists a point $\xi \in]x_1, x_2[$ such that*

$$\frac{g(x_1)f(x_2) - g(x_2)f(x_1)}{g(x_1) - g(x_2)} = f(\xi) - \frac{g(\xi)}{g'(\xi)} f'(\xi). \quad (3.27)$$

Here it is assumed that neither $g(x)$ nor $g'(x)$ is ever zero in $[a, b]$.

Obviously ξ depends on x_1 and x_2 and one may ask for what f and g the mean value ξ depends on x_1 and x_2 in a given manner. From this point of view (3.27) becomes a functional equation. The right side of (3.27) can be replaced by an unknown function h to get

$$g(x_1)f(x_2) - g(x_2)f(x_1) = [g(x_1) - g(x_2)] h(\xi).$$

Replacing x_1 by x and x_2 by y , and assuming ξ to be the arithmetic mean of x and y , one obtains

$$g(x)f(y) - g(y)f(x) = [g(x) - g(y)] h\left(\frac{x+y}{2}\right), \quad x, y \in [a, b], \quad (3.28)$$

where $[a, b]$ is a proper interval not containing zero. Kuczma (1991) gave the solution of a similar equation and left the solution of this equation to the reader. His proof is a bit involved since he worked on a proper interval $[a, b]$. We will solve the above functional equation following his method but we have replaced the proper interval $[a, b]$ by the whole real line \mathbb{R} . Further,

we also assume that $g(\kappa) = 0$ for some $\kappa \in \mathbb{R}$. This is done to make the proof less technical and easily readable.

Theorem 3.5 *Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous and strictly increasing function with $g(\kappa) = 0$ for some $\kappa \in \mathbb{R}$. The functions $f, g, h : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$g(x)f(y) - g(y)f(x) = [g(x) - g(y)] h\left(\frac{x+y}{2}\right), \quad x, y \in \mathbb{R}. \quad (3.29)$$

if and only if

$$\left. \begin{array}{l} f(x) = \alpha g(x) + \beta \\ h(x) = \beta \\ g(x) = \text{arbitrary,} \end{array} \right\} \quad (3.30)$$

where α, β are arbitrary constants.

Proof: Since g is strictly increasing, we note that $g \neq 0$. Thus, there exists a positive δ in \mathbb{R} such that $g(\delta) = k$, where k is a nonzero constant. Note that if g satisfies (3.29) so also cg , where c is an arbitrary constant. Hence, we assume without loss of generality $g(\delta) = 1$. Further, we observe that the functions

$$f_\delta(x) = f(\delta x), \quad g_\delta(x) = g(\delta x), \quad \text{and} \quad h_\delta(x) = h(\delta x) \quad (3.31)$$

satisfy the functional equation (3.29) if f, g and h do. That is

$$g_\delta(x)f_\delta(y) - g_\delta(y)f_\delta(x) = [g_\delta(x) - g_\delta(y)] h_\delta\left(\frac{x+y}{2}\right), \quad (3.32)$$

for all $x, y \in \mathbb{R}$. Moreover, we see that $g_\delta(1) = 1$. Since g is continuous and strictly increasing so also g_δ . For arbitrary constants α and β , if we define

$$\left. \begin{array}{l} F(x) = f_\delta(x) - \alpha g_\delta(x) - \beta \\ H(x) = h_\delta(x) - \beta \\ G(x) = g_\delta(x), \end{array} \right\} \quad (3.33)$$

then from (3.32) and (3.33), we get the following equation

$$F(x)G(y) - F(y)G(x) = [G(y) - G(x)]H\left(\frac{x+y}{2}\right). \quad (3.34)$$

If we choose $\alpha = f_\delta(1) - h_\delta(1)$ and $\beta = h_\delta(1)$, then from (3.33) we see that

$$F(1) = 0, \quad H(1) = 0, \quad \text{and} \quad G(1) = 1. \quad (3.35)$$

Letting $y = 1$ in (3.34) and using (3.35), we get

$$F(x) = [1 - G(x)] H\left(\frac{1+x}{2}\right). \quad (3.36)$$

Substituting (3.36) in (3.34), we see that

$$\begin{aligned} & G(y)[1 - G(x)] H\left(\frac{1+x}{2}\right) - G(x)[1 - G(y)] H\left(\frac{1+y}{2}\right) \\ &= [G(y) - G(x)] H\left(\frac{x+y}{2}\right) \end{aligned} \quad (3.37)$$

for all $x, y \in \mathbb{R}$. Since g is continuous and monotonic so also G (see (3.31) and (3.33)). Hence, continuity and monotonicity of G imply that

$$G(x) \neq 0 \quad \text{for all } x \in \mathbb{R} \setminus \{x_o\} \quad (3.38)$$

for some $x_o \in \mathbb{R}$. Dividing (3.37) by $G(x)G(y)$, we obtain

$$\begin{aligned} & \frac{1 - G(x)}{G(x)} H\left(\frac{1+x}{2}\right) - \frac{1 - G(y)}{G(y)} H\left(\frac{1+y}{2}\right) \\ &= \frac{G(y) - G(x)}{G(x)G(y)} H\left(\frac{x+y}{2}\right) \end{aligned} \quad (3.39)$$

for all $x, y \in \mathbb{R} \setminus \{x_o\}$. Substituting

$$z = \frac{1}{2}(x+1) \quad \text{and} \quad w = \frac{1}{2}(y+1) \quad (3.40)$$

in (3.39), we obtain

$$\begin{aligned} & \frac{1 - G(2z-1)}{G(2z-1)} H(z) - \frac{1 - G(2w-1)}{G(2w-1)} H(w) \\ &= -\frac{G(2z-1) - G(2w-1)}{G(2z-1)G(2w-1)} H(z+w-1) \end{aligned} \quad (3.41)$$

for all $z, w \in \mathbb{R} \setminus \{\frac{1}{2}(x_o+1)\}$. Putting $z = 2-w$ in (3.41) and using (3.35), we get

$$\frac{G(3-2w)-1}{G(3-2w)} H(2-w) = \frac{G(2w-1)-1}{G(2w-1)} H(w) \quad (3.42)$$

for all $w \in \mathbb{R} \setminus \{\frac{1}{2}(x_o + 1), \frac{1}{2}(3 - x_o)\}$. Next, we replace w by $2 - w$ in (3.41) to obtain

$$\begin{aligned} & \frac{G(2z - 1) - 1}{G(2z - 1)} H(z) - \frac{G(3 - 2w) - 1}{G(3 - 2w)} H(2 - w) \\ &= \frac{G(2z - 1) - G(3 - 2w)}{G(2z - 1)G(3 - 2w)} H(z - w + 1) \end{aligned} \quad (3.43)$$

for $z \in \mathbb{R} \setminus \{\frac{1}{2}(x_o + 1)\}$ and $w \in \mathbb{R} \setminus \{\frac{1}{2}(x_o + 1), \frac{1}{2}(3 - x_o)\}$. Now using (3.42) in (3.43), we see that

$$\begin{aligned} & \frac{G(2z - 1) - 1}{G(2z - 1)} H(z) - \frac{G(2w - 1) - 1}{G(2w - 1)} H(w) \\ &= \frac{G(2z - 1) - G(3 - 2w)}{G(2z - 1)G(3 - 2w)} H(z - w + 1) \end{aligned} \quad (3.44)$$

for $z \in \mathbb{R} \setminus \{\frac{1}{2}(x_o + 1)\}$ and $w \in \mathbb{R} \setminus \{\frac{1}{2}(x_o + 1), \frac{1}{2}(3 - x_o)\}$. Comparing (3.44) with (3.41), we get

$$\begin{aligned} & \frac{G(2z - 1) - G(2w - 1)}{G(2z - 1)G(2w - 1)} H(z + w - 1) \\ &= \frac{G(2z - 1) - G(3 - 2w)}{G(2z - 1)G(3 - 2w)} H(z - w + 1) \end{aligned}$$

for $z \in \mathbb{R} \setminus \{\frac{1}{2}(x_o + 1)\}$ and $w \in \mathbb{R} \setminus \{\frac{1}{2}(x_o + 1), \frac{1}{2}(3 - x_o)\}$. By (3.38), we see that $G(2z - 1) \neq 0$ for $z \in \mathbb{R} \setminus \{\frac{1}{2}(x_o + 1)\}$. Hence the above equation reduces to

$$\begin{aligned} & \frac{G(2z - 1) - G(2w - 1)}{G(2w - 1)} H(z + w - 1) \\ &= \frac{G(2z - 1) - G(3 - 2w)}{G(3 - 2w)} H(z - w + 1). \end{aligned} \quad (3.45)$$

Introducing new variables

$$t = z + w - 2 \quad \text{and} \quad s = z - w \quad (3.46)$$

so that

$$z = 1 + \frac{1}{2}(t + s) \quad \text{and} \quad w = 1 + \frac{1}{2}(t - s). \quad (3.47)$$

Thus using (3.46) in (3.45), we get

$$\begin{aligned} & \frac{G(1+t+s) - G(1+t-s)}{G(1+t-s)} H(1+t) \\ &= \frac{G(1+t+s) - G(1-t+s)}{G(1-t+s)} H(1+s) \end{aligned} \quad (3.48)$$

for $t, s \in \mathbb{R}$ with $t - s \neq \pm(x_o - 1)$. If we define

$$w(t) = H(1+t) \quad \text{for } t \in \mathbb{R} \quad (3.49)$$

and

$$z(t, s) = \frac{G(1+s+t) - G(1+s-t)}{G(1+s+t) - G(1+s-t)} \frac{G(1+t-s)}{G(1+t-s)} \quad (3.50)$$

for all $t, s \in \mathbb{R}$ with $t - s \neq \pm(x_o - 1)$ and $s \neq 0$. Since g is continuous and strictly increasing so also G . Note that in (3.50)

$$G(1+t+s) - G(1+t-s) \neq 0$$

for $s \neq 0$ because of the strict monotonicity of G . Further, $G(1+s-t) \neq 0$ for all $t, s \in \mathbb{R}$ with $t - s \neq \pm(x_o - 1)$ since $G(x) \neq 0$ for all $x \in \mathbb{R} \setminus \{x_o\}$ by the continuity of G . Thus the definition of $z(t, s)$ in (3.50) makes sense.

Then the use of (3.49) and (3.50) in (3.48) yields

$$w(t) = z(t, s) w(s) \quad (3.51)$$

for $t, s \in \mathbb{R}$ with $t - s \neq \pm(x_o - 1)$ and $s \neq 0$. From (3.51) one obtains either

$$w(t) \neq 0 \quad \text{for all } t \in \mathbb{R} \setminus \{0\} \quad (3.52)$$

or

$$w(t) = 0 \quad \text{for all } t \in \mathbb{R}.$$

Now we consider two cases.

Case 1. First, we treat the case when $w(t) \neq 0$ for $t \in \mathbb{R} \setminus \{0\}$. We fix $s = \xi \in \mathbb{R} \setminus \{0\}$ and obtain from (3.51)

$$w(t) = c_o z(t, \xi), \quad \text{for all } t \in \mathbb{R} \setminus \{0\}, \quad (3.53)$$

where $c_0 = w(\xi)$ is a real constant. Substituting (3.53) in (3.51), we obtain

$$z(t, \xi) = z(t, s) z(s, \xi), \quad t, s, \xi \in \mathbb{R} \setminus \{0\}. \quad (3.54)$$

Since we have unrestricted choice of ξ in $\mathbb{R} \setminus \{0\}$, we can consider ξ as an independent variable. The functional equation (3.54) is the well known Sincov equation (see Aczél (1966)) and hence we have

$$z(t, s) = \frac{\phi(t+1)}{\phi(s+1)} \quad \text{for all } t, s \in \mathbb{R}, s \neq -1, \quad (3.55)$$

where $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is an arbitrary function satisfying

$$\phi(t) \neq 0 \quad t \in \mathbb{R} \setminus \{-1\} \quad \text{with } \phi(-1) = 0. \quad (3.56)$$

By (3.55) and (3.53), we observe that

$$w(t) = c \phi(t+1), \quad t \in \mathbb{R}, \quad (3.57)$$

where c is a real constant. Therefore, by (3.57) and (3.49), we obtain

$$H(t) = c \phi(t) \quad t \in \mathbb{R}. \quad (3.58)$$

This with (3.36) and (3.33) yields

$$\begin{aligned} f(x) &= \alpha g(x) + \gamma [1 - g(x)] \phi\left(\frac{1+x}{2}\right) + \beta \\ h(x) &= \gamma \phi(x) + \beta \\ g(x) &= \text{arbitrary,} \end{aligned}$$

where α, β, γ are arbitrary constants, and $\phi : \mathbb{R} \rightarrow \mathbb{R}$ an arbitrary function satisfying (3.56). Letting the above form of f, g, h into (3.29), we get

$$\begin{aligned} g(y) [1 - g(x)] \phi\left(\frac{1+x}{2}\right) - g(x) [1 - g(y)] \phi\left(\frac{1+y}{2}\right) \\ = [g(y) - g(x)] \phi\left(\frac{x+y}{2}\right) \end{aligned} \quad (3.59)$$

Recall that by assumption $g(\kappa) = 0$ for $\kappa \in \mathbb{R}$. Substituting $x = \kappa$ in (3.59), we get

$$g(y) \phi\left(\frac{1+\kappa}{2}\right) = g(y) \phi\left(\frac{\kappa+y}{2}\right).$$

Thus, if $y \neq \kappa$, we have

$$\phi\left(\frac{1+\kappa}{2}\right) = \phi\left(\frac{\kappa+y}{2}\right),$$

that is $\phi(y) = \gamma_o$, where γ_o is an arbitrary constant. However, letting $x = \kappa - 1$ and $y = \kappa + 1$ in (3.59), we see that

$$\gamma_o [g(\gamma+1) - g(\gamma-1)] = [g(\gamma+1) - g(\gamma-1)] \phi(\kappa).$$

Thus, $\phi(t) = \gamma_o$ for all $t \in \mathbb{R}$. Hence we have the asserted solution.

Case 2. Next, we consider the case when $w(t) = 0$ for all $t \in \mathbb{R}$. Hence from (3.49) we have $H(t) = 0$ for all $t \in \mathbb{R}$. From (3.36) and the fact that G is continuous with $G(1) = 1$ we have $f(x) = \alpha g(x)$. Thus, we get the asserted solution with $\beta = 0$. Since no more cases are left, the proof of the theorem is now complete.

3.4 Equations Motivated by Simpson's Rule

Simpson's rule is an elementary numerical method for evaluating a definite integral $\int_a^b f(t) dt$. The method consists of partitioning the interval $[a, b]$ into subintervals of equal lengths and then approximating the graph of f over each subinterval with a quadratic function. If $a = x_0 < x_1 < x_2 < \dots < x_{2n} = b$ is a partition of $[a, b]$ into $2n$ subintervals, each of length $\frac{b-a}{2n}$, then

$$\begin{aligned} & \int_a^b f(t) dt \\ & \simeq \frac{b-a}{6n} [f(x_0) + 4f(x_1) + \dots + 2f(x_{2n-2}) + 4f(x_{2n-1}) + f(x_{2n})]. \end{aligned}$$

This approximation formula is called Simpson's rule. It is well known that the error bound for Simpson's rule approximation is

$$\begin{aligned} & \left| \int_a^b f(t) dt - \frac{b-a}{6n} [f(x_0) + \dots + 2f(x_{2n-2}) + 4f(x_{2n-1}) + f(x_{2n})] \right| \\ & \leq \frac{K(b-a)^5}{180 n^4} \end{aligned} \tag{3.60}$$

where $K = \sup \{ |f^{(4)}(x)| \mid x \in [a, b] \}$. It is easy to note from this inequality that if f is four times continuously differentiable and $f^{(4)}(x) = 0$, then

$$\begin{aligned} & \int_a^b f(t) dt \\ &= \frac{b-a}{6n} [f(x_0) + 4f(x_1) + \cdots + 2f(x_{2n-2}) + 4f(x_{2n-1}) + f(x_{2n})]. \end{aligned}$$

This is obviously true if $n = 1$ and it reduces to

$$\int_a^b f(t) dt = \frac{b-a}{6} [f(x_0) + 4f(x_1) + f(x_2)].$$

Letting $a = x$, $b = y$ and $x_1 = \frac{x+y}{2}$ in the above formula, we obtain

$$\int_x^y f(t) dt = \frac{y-x}{6} \left[f(x) + 4f\left(\frac{x+y}{2}\right) + f(y) \right]. \quad (3.61)$$

This integral equation (3.61) holds for all $x, y \in \mathbb{R}$ if f is a polynomial of degree at most three. However, it is not obvious that if (3.61) holds for all $x, y \in \mathbb{R}$, then the only solution f is a cubic polynomial. The integral equation (3.61) leads to the functional equation

$$g(y) - g(x) = \frac{y-x}{6} \left[f(x) + 4f\left(\frac{x+y}{2}\right) + f(y) \right] \quad (3.62)$$

where g is an antiderivative of f . The above equation is a special case of the functional equation

$$f(x) - g(y) = (x-y)[h(x+y) + \psi(x) + \phi(y)] \quad (3.63)$$

for all $x, y \in \mathbb{R}$. In this section, we determine the general solution of the above functional equation (3.63). The following two functional equations

$$g(x) - g(y) = (x-y)f(x+y) + (x+y)f(x-y), \quad (3.64)$$

and

$$xf(y) - yf(x) = (x-y)[g(x+y) - g(x) - g(y)], \quad (3.65)$$

are instrumental in solving the functional equation (3.63). The functional equation (3.64) can be considered as a variant of

$$\frac{f(x) - f(y)}{x - y} = h(x+y)$$

and obtained by adding an extra term $(x + y)h(x - y)$ to the equation

$$f(x) - f(y) = (x - y)h(x + y).$$

The equation (3.65) is another variant of

$$xf(y) - yf(x) = (x - y)g(x + y)$$

and it is obtained by replacing the term $g(x + y)$ by the Cauchy difference of g , that is, $g(x + y) - g(x) - g(y)$.

The following theorem is needed to establish the next result.

Theorem 3.6 *The functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$g(x) - g(y) = (x - y)f(x + y) + (x + y)f(x - y), \quad \forall x, y \in \mathbb{R} \quad (3.66)$$

if and only if

$$f(x) = ax^3 + A(x) \quad \text{and} \quad g(x) = 2ax^4 + 2xA(x) + b, \quad (3.67)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is additive, a and b are arbitrary constants.

Proof: Letting $y = 0$ in (3.66), we see that

$$g(x) = 2xf(x) + b, \quad (3.68)$$

where $b = g(0)$. Substitution of (3.68) into (3.66) yields

$$2xf(x) - 2yf(y) = (x - y)f(x + y) + (x + y)f(x - y). \quad (3.69)$$

Letting $x = y$ in (3.66), we get $f(0) = 0$. Next, we let $x = 0$ to obtain

$$g(y) = yf(y) - yf(-y) + b.$$

This with (3.68) yields $f(-y) = -f(y)$ for $y \neq 0$. In (3.69), we replace y by $y + z$ and obtain

$$\begin{aligned} & 2xf(x) - 2(y + z)f(y + z) \\ &= (x - y - z)f(x + y + z) + (x + y + z)f(x - y - z). \end{aligned} \quad (3.70)$$

Similarly, replacing y by z and x by $x + y$ in (3.69), we get

$$\begin{aligned} & 2(x + y)f(x + y) - 2zf(z) \\ &= (x + y - z)f(x + y + z) + (x + y + z)f(x + y - z). \end{aligned} \quad (3.71)$$

Adding equations (3.70) and (3.71), one obtains

$$\begin{aligned} 2xf(x) - 2zf(z) + 2(x+y)f(x+y) - 2(y+z)f(y+z) \\ = 2(x-z)f(x+y+z) \\ + (x+y+z)[f(x-y-z) + f(x+y-z)]. \end{aligned} \quad (3.72)$$

Using (3.69) twice to replace the terms on the left side of (3.72), we get

$$\begin{aligned} (x-z)f(x+z) + (x+z)f(x-z) + (x-z)f(x+2y+z) + (x+2y+z)f(x-z) \\ = 2(x-z)f(x+y+z) \\ + (x+y+z)[f(x-y-z) + f(x+y-z)]. \end{aligned} \quad (3.73)$$

Now, letting $z = -x$ in (3.73), we see that

$$2xf(2y) + 2yf(2x) = 4xf(y) + y[f(2x-y) + f(2x+y)] \quad (3.74)$$

Substitution of $u = 2x$ in (3.74) yields

$$\frac{u}{y} f(2y) + 2f(u) - \frac{2u}{y} f(y) = f(u-y) + f(u+y) \quad (3.75)$$

for $y \neq 0$. Interchanging u and y (so now $u \neq 0$ as well) and using that f is odd, we have

$$\frac{y}{u} f(2u) + 2f(y) - \frac{2y}{u} f(u) = f(u+y) - f(u-y). \quad (3.76)$$

Adding (3.75) and (3.76), we obtain

$$\begin{aligned} f(u+y) - f(u) - f(y) \\ = \frac{u}{2y}[f(2y) - 2f(y)] + \frac{y}{2u}[f(2u) - 2f(u)] \end{aligned} \quad (3.77)$$

for all $u, y \in \mathbb{R} \setminus \{0\}$. Define

$$h(x) = \frac{f(2x) - 2f(x)}{2x} \quad \text{for } x \neq 0. \quad (3.78)$$

Note that h is even since f is odd. Use of (3.78) in (3.77) yields

$$f(u+y) - f(u) - f(y) = uh(y) + yh(u) \quad (3.79)$$

where $u, y \in \mathbb{R} \setminus \{0\}$. Let

$$H(u, v) = f(u+v) - f(u) - f(v) \quad (3.80)$$

be the Cauchy difference of f . Hence H satisfies

$$H(u+v, w) + H(u, v) = H(u, v+w) + H(v, w) \quad (3.81)$$

for all $u, v, w \in \mathbb{R}$. From (3.80), (3.79) and (3.81), after some simplifications, we see that

$$w [h(u+v) - h(u) - h(v)] = u [h(v+w) - h(v) - h(w)] \quad (3.82)$$

for $u, v, w, u+v, v+w \in \mathbb{R} \setminus \{0\}$. Letting $w = v$ in (3.82), we get

$$v [h(u+v) - h(u) - h(v)] = u [h(2v) - 2h(v)].$$

Interchanging u and v , we obtain

$$u [h(u+v) - h(u) - h(v)] = v [h(2u) - 2h(u)]. \quad (3.83)$$

Thus, we have

$$u^2 [h(2v) - 2h(v)] = v^2 [h(2u) - 2h(u)].$$

Hence

$$h(2u) - 2h(u) = 6au^2 \quad \text{for all } u \neq 0, \quad (3.84)$$

where a is a constant. Using (3.84) in (3.83), we have

$$h(u+v) - h(u) - h(v) = 6auv, \quad u, v \in \mathbb{R} \setminus \{0\}$$

which can be rearranged as

$$h(u+v) - 3a(u+v)^2 = h(u) - 3au^2 + h(v) - 3av^2. \quad (3.85)$$

Letting

$$A_o(u) = h(u) - 3au^2, \quad (3.86)$$

in (3.85), we get

$$A_o(u+v) = A_o(u) + A_o(v) \quad (3.87)$$

for all $u, v, u+v \in \mathbb{R} \setminus \{0\}$. Since

$$A_o(1) = A_o(u+1-u) = A_o(u) + A_o(1) + A_o(-u),$$

we get $A_o(-u) = -A_o(u)$ for $u \neq 0, 1$. But $A_o(2) = 2A_o(1)$ and $A_o(-2) = 2A_o(-1)$, that is $A_o(-1) = -A_o(1)$. Hence $A_o(-u) = -A_o(u)$ for all $u \neq 0$. Changing u to $-u$ in (3.86) and using the fact that h is even, we obtain

$$h(x) = 3ax^2 \quad \text{for all } x \neq 0. \quad (3.88)$$

Letting (3.88) into (3.79), we obtain

$$f(u+v) - f(u) - f(v) = 3au^2v + 3auv^2 \quad (3.89)$$

for all $u, v \in \mathbb{R} \setminus \{0\}$. This in turn gives a Cauchy equation (see Aczél and Dhombres (1988))

$$f(u+v) - a(u+v)^3 = f(u) - au^3 + f(v) - av^3,$$

and hence

$$f(u) = au^3 + A(u)$$

for all $u \in \mathbb{R}$ since $f(0) = 0$. Here $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function. From (3.68), we get the form of g as asserted in (3.67) and now the proof is complete.

Now we present the general solution of the functional equation (3.90) without any regularity assumption on the unknown functions.

Theorem 3.7 *The functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$xf(y) - yf(x) = (x-y)[g(x+y) - g(x) - g(y)] \quad (3.90)$$

for all $x, y \in \mathbb{R}$ if and only if

$$\left. \begin{aligned} f(x) &= 3ax^3 + 2bx^2 + cx + d \\ g(x) &= -ax^3 - bx^2 - A(x) - d, \end{aligned} \right\} \quad (3.91)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is additive and a, b, c, d are arbitrary constants.

Proof: Letting $x = 0$ in (3.90) yields $yf(0) = -yg(0)$ for all $y \in \mathbb{R}$. Choosing $y \neq 0$, we see that $f(0) = -g(0)$. Substituting $y = -x$ in (3.90), we obtain

$$x[f(x) + f(-x)] = 2x[g(0) - g(x) - g(-x)]$$

for all $x \in \mathbb{R}$. Hence

$$f(x) + f(-x) = 2[g(0) - g(x) - g(-x)] \quad (3.92)$$

for all $x \in \mathbb{R} \setminus \{0\}$. But in view of $f(0) = -g(0)$, we see that (3.92) holds for all $x \in \mathbb{R}$. Next, replacing x by $-x$ in (3.90), we get

$$yf(-x) + xf(y) = (y+x)[g(y-x) - g(-x) - g(y)]. \quad (3.93)$$

Subtracting (3.90) from (3.93), we have

$$\begin{aligned} y[f(x) + f(-x)] &= y[g(x+y) + g(y-x) - g(x) - g(-x) - 2g(y)] \\ &\quad + x[g(y-x) - g(x+y) + g(x) - g(-x)] \end{aligned} \quad (3.94)$$

for all $x, y \in \mathbb{R}$. Using (3.92) in (3.94), we obtain

$$\begin{aligned} 2y[g(0) - g(x) - g(-x)] &= y[g(x+y) + g(y-x) - g(x) - g(-x) - 2g(y)] \\ &\quad + x[g(y-x) - g(x+y) + g(x) - g(-x)] \end{aligned}$$

that is

$$\begin{aligned} y[g(x+y) + g(y-x) + g(x) + g(-x) - 2g(y) - 2g(0)] \\ = x[g(y+x) - g(y-x) - g(x) + g(-x)]. \end{aligned} \quad (3.95)$$

If $g(x)$ satisfies (3.95), so does $g(x) - g(0)$. Thus, without loss of generality we may suppose that $g(0) = 0$. Then (3.95) reduces to

$$\begin{aligned} y[g(x+y) + g(y-x) + g(x) + g(-x) - 2g(y)] \\ = x[g(y+x) - g(y-x) - g(x) + g(-x)], \end{aligned}$$

which is

$$\begin{aligned} (y-x)g(x+y) + (x+y)g(y-x) + y[g(x) + g(-x)] \\ = 2yg(y) - x[g(x) - g(-x)]. \end{aligned} \quad (3.96)$$

Replacing y with $-y$ in (3.96), we obtain

$$\begin{aligned} -(y+x)g(x-y) + (x-y)g(-y-x) \\ = y[g(x) + g(-x)] - 2yg(-y) - x[g(x) - g(-x)] = 0. \end{aligned} \quad (3.97)$$

Subtracting (3.97) from (3.96), we have

$$\begin{aligned} (y+x)[g(y-x) + g(x-y)] + (y-x)[g(x+y) + g(-y-x)] \\ = 2y[g(y) + g(-y)] - 2y[g(x) + g(-x)]. \end{aligned} \quad (3.98)$$

Defining

$$\phi(x) = g(x) + g(-x) \quad (3.99)$$

we see that ϕ is even and using it in (3.98), we obtain

$$(y+x)\phi(y-x) + (y-x)\phi(x+y) = 2y\phi(y) - 2y\phi(x). \quad (3.100)$$

Interchanging x and y in (3.100) and using the fact that ϕ is even, we get

$$(y+x)\phi(y-x) - (y-x)\phi(x+y) = 2x[\phi(x) - \phi(y)]. \quad (3.101)$$

Adding (3.100) and (3.101), we have

$$(y+x)\phi(y-x) = (x-y)[\phi(x) - \phi(y)]. \quad (3.102)$$

Letting $2u = x + y$ and $2v = x - y$ in (3.102), we get

$$u\phi(2v) = v[\phi(u+v) - \phi(u-v)].$$

Again interchanging u with v in the above and using the fact that ϕ is even, we get

$$v\phi(2u) = u[\phi(u+v) - \phi(u-v)].$$

Hence from the above two equations, we see that

$$u^2\phi(2v) = v^2\phi(2u)$$

that is $\phi(u) = -2bu^2$, where b is a constant. Therefore, in view of (3.99), we get

$$g(x) + g(-x) = -2bx^2. \quad (3.103)$$

Now adding (3.97) to (3.96), we get

$$\begin{aligned} & (y+x)[g(y-x) - g(x-y)] + (y-x)[g(x+y) - g(-x-y)] \\ &= 2y[g(y) - g(-y)] - 2x[g(x) - g(-x)]. \end{aligned} \quad (3.104)$$

Defining

$$\psi(x) = g(x) - g(-x) \quad (3.105)$$

we observe that ψ is odd and using it in (3.104), we see that

$$2x\psi(x) - 2y\psi(y) = (x-y)\psi(x+y) + (x+y)\psi(x-y) \quad (3.106)$$

for all $x, y \in \mathbb{R}$. The general solution of (3.106) can be obtained from Theorem 3.6. Hence, we have

$$\psi(x) = -2ax^3 - 2A(x) \quad (3.107)$$

where A is additive and a is an arbitrary constant. Using (3.107), (3.105) and (3.103), we get

$$g(x) = -ax^3 - bx^2 - A(x).$$

Now removing the assumption $g(0) = 0$, we obtain

$$g(x) = -ax^3 - bx^2 - A(x) - d, \quad (3.108)$$

where d is an arbitrary constant. Letting (3.108) into (3.90) and simplifying, we have

$$y[f(x) - 3ax^3 - 2bx^2 - d] = x[f(y) - 3ay^3 - 2by^2 - d]$$

for all $x, y \in \mathbb{R}$. Hence $f(x) - 3ax^3 - 2bx^2 - d = cx$ for $x \neq 0$. Since $f(0) = -g(0)$, we get

$$f(x) = 3ax^3 + 2bx^2 + cx + d,$$

for all $x \in \mathbb{R}$. The proof of the theorem is now complete.

The following theorem is established using Theorem 3.7.

Theorem 3.8 *The functions $f, g, h, k : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$f(x) - g(y) = (x - y)[h(x + y) + k(x) + k(y)] \quad (3.109)$$

for all $x, y \in \mathbb{R}$ if and only if

$$\left. \begin{array}{lcl} f(x) & = & 3ax^4 + 2bx^3 + cx^2 + dx + \alpha \\ g(y) & = & 3ay^4 + 2by^3 + cy^2 + dy + \alpha \\ h(x) & = & ax^3 + bx^2 + A(x) + d - 2\beta \\ k(x) & = & 2ax^3 + bx^2 + cx - A(x) + \beta, \end{array} \right\} \quad (3.110)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function and $a, b, c, d, \alpha, \beta$ are arbitrary constants.

Proof: Letting $x = y$ in (3.109), we see that $f(x) = g(x)$. Hence (3.109) reduces to

$$f(x) - f(y) = (x - y)[h(x + y) + k(x) + k(y)]. \quad (3.111)$$

Putting $y = 0$ in (3.111), we obtain

$$f(x) = f(0) + x[h(x) + k(x) + k(0)]. \quad (3.112)$$

Letting (3.112) into (3.111) and rearranging, we obtain

$$\begin{aligned} & y[h(x) + k(x)] - x[h(y) + k(y)] \\ &= (x - y)[h(x + y) - h(x) - h(y) - k(0)]. \end{aligned} \quad (3.113)$$

Defining

$$\phi(x) = h(x) + k(x) \quad \text{and} \quad g(x) = -h(x) - k(0), \quad (3.114)$$

and using (3.114) in (3.113), we have

$$x\phi(y) - y\phi(x) = (x - y)[g(x + y) - g(x) - g(y)] \quad (3.115)$$

for all $x, y \in \mathbb{R}$. The general solution of (3.115) can be obtained from Theorem 3.7. Therefore

$$\left. \begin{aligned} \phi(x) &= 3ax^3 + 2bx^2 + cx + d_o \\ g(x) &= -ax^3 - bx^2 - A(x) - d_o, \end{aligned} \right\} \quad (3.116)$$

where a, b, c, d_o are constants. From (3.114) and (3.116), we obtain

$$k(x) = \phi(x) + g(x) + \beta, \quad (3.117)$$

where $\beta = k(0)$. Now using (3.116) and (3.117), we obtain

$$k(x) = 2ax^2 + bx^2 + cx - A(x) + \beta. \quad (3.118)$$

Again from (3.114) and (3.118), we have

$$h(x) = ax^3 + bx^2 + A(x) + d - 2\beta, \quad (3.119)$$

where $d = d_o + \beta$. Using (3.112), (3.114) and (3.116), we get

$$f(x) = 3ax^4 + 2bx^3 + cx^2 + dx + \alpha, \quad (3.120)$$

where $\alpha = f(0)$. The proof of the theorem is now complete.

Remark. It follows easily from Theorem 3.8 that the solution of (3.62) is $g(x) = 3ax^4 + 2bx^3 + cx^2 + dx + \alpha$ and $f(x) = 12ax^3 + 6bx^2 + 2cx + d$ as predicted. Note that the general solution is obtained without any regularity assumptions on f and g .

Theorem 3.9 *The functions $f, g, h, \phi, \psi : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$f(x) - g(y) = (x - y)[h(x + y) + \phi(x) + \psi(y)] \quad (3.121)$$

for all $x, y \in \mathbb{R}$ if and only if

$$\left. \begin{aligned} f(x) &= 3ax^4 + 2bx^3 + cx^2 + dx + \alpha \\ g(y) &= 3ay^4 + 2by^3 + cy^2 + dy + \alpha \\ h(x) &= ax^3 + bx^2 + A(x) + d - 2\beta \\ \phi(x) &= 2ax^3 + bx^2 + cx - A(x) + \beta + \gamma \\ \psi(y) &= 2ay^3 + by^2 + cy - A(y) + \beta - \gamma, \end{aligned} \right\} \quad (3.122)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function and $a, b, c, d, \alpha, \beta, \gamma$ are arbitrary constants.

Proof: First, letting $x = y$ in (3.121), we see that $f = g$. Interchanging x with y in (3.121) and using the fact $f = g$, we get

$$f(y) - f(x) = (y - x)[h(x + y) + \phi(y) + \psi(x)]. \quad (3.123)$$

Adding (3.123) to (3.121) and using $f = g$, we get $\psi(x) - \phi(x) = \psi(y) - \phi(y)$ for all $x, y \in \mathbb{R}$. Thus

$$\psi(x) = \phi(x) - 2\gamma, \quad (3.124)$$

where γ is an arbitrary constant. Putting (3.124) into (3.121), we have

$$f(x) - g(y) = (x - y)[h(x + y) + \phi(x) + \phi(y) - 2\gamma]. \quad (3.125)$$

From Theorem 3.9 and (3.124), we have the asserted solution (3.122). This completes the proof.

3.5 Some Generalizations

In this section, we consider some generalizations of the functional equation (3.62). The middle term of the equation (3.62), that is $4f(\frac{x+y}{2})$ is due

to the fact that in Simpson's rule one partitions the interval into subintervals of equal lengths. However, there is no reason why one should be restricted to such an equal partition. If we allow unequal partition, then the middle term is no longer of the form $4f\left(\frac{x+y}{2}\right)$ but rather it is of the form $\alpha f(sx + ty)$, where α, s, t are constants. Taking this into account we have a generalization of (3.62) as

$$f(x) - f(y) = (x - y)[h(sx + ty) + g(x) + g(y)] \quad (3.126)$$

for all $x, y \in \mathbb{R}$ with s and t being *a priori* chosen parameters.

Next, we determine the general solution of the functional equation (3.126) without any regularity assumptions (differentiability, continuity, measurability, etc.) imposed on the unknown functions. Further, utilizing the solution of this equation we determine the general solution of the functional equation

$$f(x) - g(y) = (x - y)[h(sx + ty) + \psi(x) + \phi(y)] \quad (3.127)$$

for all $x, y \in \mathbb{R}$ with s and t being *a priori* chosen parameters.

Note that the functional equation (3.127) includes the functional equation

$$f(x) - g(y) = (x - y) h(x + y)$$

which was studied by Haruki (1979) and Aczél (1985), and the functional equation

$$f(x) - g(y) = (x - y) h(sx + ty)$$

whose solution was determined by Kannappan, Sahoo and Jacobson (1995).

The following result is due to Haruki (1979) and will be used in determining the general solution of the functional equation (3.126).

Lemma 3.2 . *The functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$f(x) - f(y) = (x - y) \left(\frac{g(x) + g(y)}{2} \right) \quad (3.128)$$

for all $x, y \in \mathbb{R}$ if and only if

$$f(x) = ax^2 + bx + c \quad \text{and} \quad g(x) = 2ax + b \quad (3.129)$$

where a, b and c are arbitrary real constants.

Proof: Letting $y = 0$ in (3.128), we get

$$2f(x) = 2f(0) + x[g(x) + g(0)]. \quad (3.130)$$

Inserting (3.130) into (3.128), we obtain

$$x[g(y) - g(0)] = y[g(x) - g(0)].$$

Therefore

$$g(x) = 2ax + b, \quad x \neq 0, \quad (3.131)$$

where a and b are constants. Note that (3.131) holds even for $x = 0$, since $b = g(0)$. Letting (3.131) in (3.130), we have

$$2f(x) = ax^2 + bx + c, \quad (3.132)$$

where $c = f(0)$. Hence we have the asserted solution and the proof is now complete.

Now we proceed to find the general solution of (3.126) with no regularity assumptions imposed on f, g and h .

Theorem 3.10 . *Let s and t be real parameters. The real valued functions $f, g, h : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation (3.126) for all $x, y \in \mathbb{R}$ if and only if*

$$f(x) = \begin{cases} ax^2 + (b+d)x + c & \text{if } s = 0 = t \\ ax^2 + (b+d)x + c & \text{if } s = 0, t \neq 0 \\ ax^2 + (b+d)x + c & \text{if } s \neq 0, t = 0 \\ 3ax^4 + 2bx^3 + cx^2 + (d+2\beta)x + \alpha & \text{if } s = t \neq 0 \\ 2ax^3 + cx^2 + 2\beta x - A(x) + \alpha & \text{if } s = -t \neq 0 \\ ax^2 + (b+d)x + c & \text{if } 0 \neq s^2 \neq t^2 \neq 0 \end{cases}$$

$$g(x) = \begin{cases} ax + \frac{b}{2} & \text{if } s = 0 = t \\ ax + \frac{b}{2} & \text{if } s = 0, t \neq 0 \\ ax + \frac{b}{2} & \text{if } s \neq 0, t = 0 \\ 2ax^3 + bx^2 + cx - A(x) + \beta & \text{if } s = t \neq 0 \\ 3ax^2 + cx + \beta, & \text{if } s = -t \neq 0 \\ ax + \frac{b}{2} & \text{if } 0 \neq s^2 \neq t^2 \neq 0 \end{cases}$$

$$h(x) = \begin{cases} \text{arbitrary with } h(0) = d & \text{if } s = 0 = t \\ d & \text{if } s = 0, t \neq 0 \\ d & \text{if } s \neq 0, t = 0 \\ a\left(\frac{x}{t}\right)^3 + b\left(\frac{x}{t}\right)^2 + A\left(\frac{x}{t}\right) + d & \text{if } s = t \neq 0 \\ -a\left(\frac{x}{t}\right)^2 - \frac{t}{x}A\left(\frac{x}{t}\right), x \neq 0 & \text{if } s = -t \neq 0, \\ d & \text{if } 0 \neq s^2 \neq t^2 \neq 0, \end{cases}$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function and $a, b, c, d, \alpha, \beta$ are arbitrary real constants.

Proof: To prove the theorem, we consider several cases depending on the parameters s and t .

Case 1. Suppose $s = 0 = t$. Then (3.126) reduces to

$$f(x) - f(y) = (x - y)[d + g(x) + g(y)], \quad (3.133)$$

where $d = h(0)$. Defining

$$F(x) = f(x) - dx \quad \text{and} \quad G(x) = 2g(x), \quad (3.134)$$

and using (3.134) in (3.133), we obtain

$$F(x) - F(y) = (x - y) \left[\frac{G(x) + G(y)}{2} \right] \quad (3.135)$$

for all $x, y \in \mathbb{R}$. The general solution of (3.135) can be obtained from Lemma 3.2 as

$$F(x) = ax^2 + bx + c \quad \text{and} \quad G(x) = 2ax + b, \quad (3.136)$$

where a, b, c are arbitrary constants. Hence from (3.134) and (3.136), we have

$$\left. \begin{aligned} f(x) &= ax^2 + (b + d)x + c \\ g(x) &= ax + \frac{b}{2} \\ h(x) &\text{arbitrary with } h(0) = d, \end{aligned} \right\} \quad (3.137)$$

where a, b, c, d are arbitrary constants.

Case 2. Suppose $s = 0$ and $t \neq 0$. (The case $s \neq 0$ and $t = 0$ can be handled in a similar manner.) Then (3.126) reduces to

$$f(x) - f(y) = (x - y)[h(ty) + g(x) + g(y)]. \quad (3.138)$$

Letting $y = 0$ in (3.138), we obtain

$$f(x) = f(0) + x [h(0) + g(x) + g(0)]. \quad (3.139)$$

Using (3.139) in (3.138), we get

$$xg(x) - yg(y) = (x - y) [h(ty) + g(x) + g(y) - g(0) - h(0)]. \quad (3.140)$$

Interchanging x and y in (3.140), we obtain

$$yg(y) - xg(x) = (y - x) [h(tx) + g(y) + g(x) - g(0) - h(0)]. \quad (3.141)$$

Adding (3.140) to (3.141), we see that

$$h(tx) = h(ty) \quad (3.142)$$

for all $x, y \in \mathbb{R}$ with $x \neq y$. Hence from (3.142), we have

$$h(x) = d \quad \text{for all } x \in \mathbb{R}, \quad (3.143)$$

where d is an arbitrary constant. Inserting (3.143) into (3.138), we obtain

$$f(x) - f(y) = (x - y) [d + g(x) + g(y)], \quad (3.144)$$

which is (3.133). Thus, by case 1, (3.143) and (3.137) we get

$$\left. \begin{array}{l} f(x) = ax^2 + (b + d)x + c \\ g(x) = ax + \frac{b}{2} \\ h(x) = d, \end{array} \right\} \quad (3.145)$$

where a, b, c, d are arbitrary constants.

Case 3. Next suppose $s \neq 0$ and $t \neq 0$. Letting $y = 0$ and $x = 0$, separately in (3.126), we obtain

$$f(x) = f(0) + x [h(sx) + g(x) + g(0)]. \quad (3.146)$$

and

$$f(y) = f(0) + y [h(ty) + g(y) + g(0)], \quad (3.147)$$

respectively. Comparing f in (3.146) and (3.147), we have

$$h(sx) = h(tx) \quad (3.148)$$

for all $x \in \mathbb{R} \setminus \{0\}$. Letting (3.146) and (3.147) into (3.126) and rearranging terms, we get

$$\begin{aligned} y[h(sx) + g(x) - g(0)] - x[h(ty) + g(y) - g(0)] \\ = (x - y)[h(sx + ty) - h(sx) - h(ty)] \end{aligned} \quad (3.149)$$

for all $x, y \in \mathbb{R}$.

Now we consider several subcases.

Subcase 3.1. Suppose $s = t$. Then (3.149) yields

$$x\phi(y) - y\phi(x) = (x - y)[\psi(x + y) - \psi(x) - \psi(y)], \quad (3.150)$$

where

$$\phi(x) = h(tx) + g(x) - g(0) \quad \text{and} \quad \psi(x) = -h(tx). \quad (3.151)$$

The solution of the functional equation (3.150) can be obtained from Theorem 3.7 as

$$\left. \begin{array}{l} \phi(x) = 3ax^3 + 2bx^2 + cx + d \\ \psi(x) = -ax^3 - bx^2 - A(x) - d, \end{array} \right\} \quad (3.152)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ and a, b, c, d are constants. From (3.152), (3.151) and (3.146), we have the asserted solution

$$\left. \begin{array}{l} f(x) = 3ax^4 + 2bx^3 + cx^2 + (d + 2\beta)x + \alpha \\ g(x) = 2ax^3 + bx^2 + cx - A(x) + \beta \\ h(x) = a\left(\frac{x}{t}\right)^3 + b\left(\frac{x}{t}\right)^2 + A\left(\frac{x}{t}\right) + d, \end{array} \right\} \quad (3.153)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive map and $a, b, c, d, \alpha, \beta$ are arbitrary constants.

Subcase 3.2. Next, suppose $s = -t$. Then from (3.148), we have $h(tx) = h(-tx)$ for all $x \in \mathbb{R} \setminus \{0\}$. That is, h is an even function in \mathbb{R} . Now with $s = -t$ and using the evenness of h , from (3.149), we have

$$\begin{aligned} y[h(tx) + g(x) - g(0)] - x[h(ty) + g(y) - g(0)] \\ = (x - y)[h(tx - ty) - h(tx) - h(ty)] \end{aligned} \quad (3.154)$$

for all $x, y \in \mathbb{R}$. Defining

$$G(x) = h(tx) + g(x) - g(0) \quad \text{and} \quad H(x) = -h(tx) \quad (3.155)$$

we have from (3.154)

$$xG(y) - yG(x) = (x - y)[H(x - y) - H(x) - H(y)]. \quad (3.156)$$

Note that H is also an even function in view of (3.155). Replacing y by $-y$ in (3.156), we get

$$xG(-y) + yG(x) = (x + y)[H(x + y) - H(x) - H(y)]. \quad (3.157)$$

Letting $x = y$ in (3.157), we get

$$G(-x) + G(x) = 2[H(2x) - 2H(x)] \quad (3.158)$$

for all $x \neq 0$. By (3.155), (3.158) holds for $x = 0$ also. Adding (3.156) and (3.157) and using (3.158), we have

$$\begin{aligned} & (x + y)H(x + y) + (x - y)H(x - y) \\ &= 2xH(x) + 2x[H(2y) - H(y)]. \end{aligned} \quad (3.159)$$

Interchanging x with y in (3.159), we get

$$\begin{aligned} & (x + y)H(x + y) + (y - x)H(x - y) \\ &= 2yH(y) + 2y[H(2x) - H(x)]. \end{aligned} \quad (3.160)$$

Adding (3.159) to (3.160) and rearranging terms, we obtain

$$\begin{aligned} & (x + y)H(x + y) - xH(x) - yH(y) \\ &= y[H(2x) - H(x)] + x[H(2y) - H(y)]. \end{aligned} \quad (3.161)$$

The equation (3.161) yields

$$\phi(x + y) - \phi(x) - \phi(y) = y\psi(x) + x\psi(y), \quad (3.162)$$

where

$$\phi(x) = xH(x) \quad \text{and} \quad \psi(x) = H(2x) - H(x). \quad (3.163)$$

Note from (3.163) since H is even, ϕ is odd and ψ is even. Replace x by $x - y$ and y by $-y$ separately in (3.162) to get

$$\phi(x) - \phi(x - y) - \phi(y) = y\psi(x - y) + (x - y)\psi(y) \quad (3.164)$$

and

$$\phi(x - y) - \phi(x) - \phi(-y) = -y\psi(x) + x\psi(-y). \quad (3.165)$$

Adding (3.165) to (3.164) and using the fact that ϕ is odd and ψ is even, we obtain

$$y[\psi(x-y) - \psi(x) - \psi(y)] = -2x\psi(y). \quad (3.166)$$

Replacing y by $-y$ in (3.166), we have

$$y[\psi(x+y) - \psi(x) - \psi(y)] = 2x\psi(y),$$

that is

$$xy[\psi(x+y) - \psi(x) - \psi(y)] = 2x^2\psi(y) \quad (3.167)$$

for $x \neq 0$. Interchanging x and y in (3.167), we have

$$xy[\psi(x+y) - \psi(x) - \psi(y)] = 2y^2\psi(x). \quad (3.168)$$

Hence, from (3.167) and (3.168), we see that

$$2x^2\psi(y) = 2y^2\psi(x)$$

for all $x, y \in \mathbb{R} \setminus \{0\}$. Thus we have

$$\psi(x) = 3ax^2 \quad \text{for all } x \in \mathbb{R} \setminus \{0\}, \quad (3.169)$$

where a is a constant. By (3.163), (3.169) holds for $x = 0$ also. Letting (3.169) into (3.162)

$$\phi(x+y) - \phi(x) - \phi(y) = 3ax^2y + 3axy^2 \quad (3.170)$$

for all $x, y \in \mathbb{R} \setminus \{0\}$. This in turns gives a Cauchy equation

$$\phi(x+y) - a(x+y)^3 = \phi(x) - ax^3 + \phi(y) - ay^3 \quad (3.171)$$

and hence

$$\phi(x) = ax^3 + A(x) \quad (3.172)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function. From (3.172) and (3.163), we get

$$xH(x) = ax^3 + A(x). \quad (3.173)$$

Using (3.173) in (3.156), we obtain

$$x \left[G(y) + \frac{A(y)}{y} - 2ay^2 \right] = y \left[G(x) + \frac{A(x)}{x} - 2ax^2 \right] \quad (3.174)$$

for all $x, y \in \mathbb{R} \setminus \{0\}$ with $x \neq y$. Thus

$$G(x) = 2ax^2 + cx - \frac{A(x)}{x}, \quad x \neq 0, \quad (3.175)$$

where c is a constant. From (3.146), (3.155), (3.173) and (3.175), we have the asserted solution

$$\left. \begin{aligned} f(x) &= 2ax^3 + cx^2 + 2\beta x - A(x) + \alpha \\ g(x) &= 3ax^2 + cx + \beta \\ h(x) &= -a \left(\frac{x}{t}\right)^2 - \frac{t}{x} A\left(\frac{x}{t}\right), \quad x \neq 0, \end{aligned} \right\} \quad (3.176)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive map and a, c, α, β are arbitrary constants.

Subcase 3.3. Suppose $s^2 \neq t^2$, that is $\det \begin{pmatrix} s & t \\ t & s \end{pmatrix} \neq 0$. Note that if x and y are linearly independent, so also $u = sx + ty$ and $v = sy + tx$. Suppose not, then for some constants a and b (not both zero), we have $0 = au + bv = (as+bt)x + (at+bs)y$. Since x and y are linearly independent, we have

$$\begin{pmatrix} s & t \\ t & s \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Since $\det \begin{pmatrix} s & t \\ t & s \end{pmatrix} \neq 0$, this yields that both a and b are zero which is a contradiction.

Now we return to equation (3.149). Using (3.148) in (3.149), we have

$$\begin{aligned} y [h(sx) + g(x) - g(0)] - x [h(sy) + g(y) - g(0)] \\ = (x - y) [h(sx + ty) - h(sx) - h(sy)] \end{aligned} \quad (3.177)$$

for all $x, y \in \mathbb{R}$. Interchanging x and y in (3.177) we have

$$\begin{aligned} x [h(sy) + g(y) - g(0)] - y [h(sx) + g(x) - g(0)] \\ = (y - x) [h(sy + tx) - h(sy) - h(sx)]. \end{aligned} \quad (3.178)$$

Adding (3.177) to (3.178), we obtain

$$h(sx + ty) = h(sy + tx) \quad (3.179)$$

for all $x, y \in \mathbb{R} \setminus \{0\}$ with $x \neq y$. Hence

$$h(x) = d, \quad x \in \mathbb{R} \setminus \{0\}, \quad (3.180)$$

where d is a constant. Using (3.180) in (3.145), we get

$$f(x) - f(y) = (x - y)[d + g(x) + g(y)]. \quad (3.181)$$

Hence, following a argument similar to case 1, we obtain the asserted solution

$$\left. \begin{array}{l} f(x) = ax^2 + (b+d)x + c \\ g(x) = ax + \frac{b}{2} \\ h(x) = d, \end{array} \right\} \quad (3.182)$$

where a, b, c, d are arbitrary constants. Since no more cases are left, the proof of the theorem is now complete.

The following result is obvious from Theorem 3.10 and it was established in Kannappan, Sahoo and Jacobson (1995) to answer a problem posed by Walter Rudin (1989). For a direct proof see Theorem 2.5 in Chapter 2.

Corollary 3.2 *Let s and t be parameters. The functions $f, g, h : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$f(x) - g(y) = (x - y)h(sx + ty)$$

for all $x, y \in \mathbb{R}$ if and only if $g(x) = f(x)$ and

$$\begin{aligned} f(x) &= \begin{cases} dx + c & \text{if } s = 0 = t \\ dx + c & \text{if } s = 0, t \neq 0 \\ dx + c & \text{if } s \neq 0, t = 0 \\ cx^2 + dx + \alpha & \text{if } s = t \neq 0 \\ \alpha - A(x) & \text{if } s = -t \neq 0 \\ dx + c & \text{if } 0 \neq s^2 \neq t^2 \neq 0 \end{cases} \\ h(x) &= \begin{cases} \text{arbitrary with } h(0) = d & \text{if } s = 0 = t \\ d & \text{if } s = 0, t \neq 0 \\ d & \text{if } s \neq 0, t = 0 \\ \frac{cx}{s} + d & \text{if } s = t \neq 0 \\ -\frac{t}{x} A\left(\frac{x}{t}\right), x \neq 0 & \text{if } s = -t \neq 0, \\ d & \text{if } 0 \neq s^2 \neq t^2 \neq 0, \end{cases} \end{aligned}$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function and c, d, α are arbitrary real constants.

The following lemma will be used to determine the general solution of the functional equation (3.127), namely

$$f(x) - g(y) = (x - y)[h(sx + ty) + \psi(x) + \phi(y)]$$

for all $x, y \in \mathbb{R}$ with s and t being *a priori* chosen parameters.

Lemma 3.3 *Let α be a nonzero real constant. The functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation*

$$f(x) - f(y) = (x - y)[\alpha xy + g(x) + g(y)] \quad (3.183)$$

for all $x, y \in \mathbb{R}$ if and only if

$$\left. \begin{array}{l} f(x) = \alpha x^3 + \beta x^2 + 2\gamma x + \delta \\ g(x) = \alpha x^2 + \beta x + \gamma, \end{array} \right\} \quad (3.184)$$

where β, γ, δ are arbitrary constants.

Proof: It is easy to check that (3.184) satisfies (3.183). To prove the converse, let $y = 0$ in (3.183) and we have

$$f(x) = \delta + x[g(x) + \gamma], \quad (3.185)$$

where $\delta = f(0)$ and $\gamma = g(0)$. Letting (3.185) in (3.183) and simplifying, we have

$$y[g(x) - \alpha x^2 - \gamma] = x[g(y) - \alpha y^2 - \gamma]$$

for all $x, y \in \mathbb{R}$. Thus

$$g(x) = \alpha x^2 + \beta x + \gamma. \quad (3.186)$$

By (3.186) and (3.185), we have

$$f(x) = \alpha x^3 + \beta x^2 + 2\gamma x + \delta \quad (3.187)$$

and the proof of the lemma is now complete.

Now we proceed to determine the general solution of the functional equation (3.127).

Theorem 3.11 *Let s and t be real parameters. The functions $f, g, h, \phi, \psi : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the functional equation (3.127) for all $x, y \in \mathbb{R}$ if and only*

if $g(x) = f(x)$ and

$$\begin{aligned}
 f(x) &= \begin{cases} ax^2 + (b+d)x + c & \text{if } s = 0 = t \\ ax^2 + bx + c & \text{if } s = 0, t \neq 0 \\ ax^2 + bx + c & \text{if } s \neq 0, t = 0 \\ 3ax^4 + 2bx^3 + cx^2 + (d+2\beta)x + \alpha & \text{if } s = t \neq 0 \\ 2ax^3 + cx^2 + (2\beta - d)x - A(x) + \alpha & \text{if } s = -t \neq 0 \\ -2bstx^3 + \beta x^2 + (2\gamma + \alpha - d)x + \delta & \text{if } 0 \neq s^2 \neq t^2 \neq 0 \end{cases} \\
 \phi(x) &= \begin{cases} ax + \frac{(b-\delta)}{2} & \text{if } s = 0 = t \\ ax + \frac{(b+\delta)}{2} & \text{if } s = 0 t \neq 0 \\ ax + \frac{(b+\delta)}{2} & \text{if } s \neq 0 t = 0 \\ 2ax^3 + bx^2 + cx - A(x) + \beta + \frac{\delta}{2} & \text{if } s = t \neq 0 \\ 3ax^2 + cx - \frac{1}{2}A_o(x) + \beta, & \text{if } s = -t \neq 0 \\ bs(s-2t)x^2 + \beta x + A(sx) + \gamma + \alpha & \text{if } 0 \neq s^2 \neq t^2 \neq 0 \end{cases} \\
 \psi(x) &= \begin{cases} ax + \frac{(b+\delta)}{2} & \text{if } s = 0 = t \\ ax + \frac{(b-\delta)}{2} - h(tx) & \text{if } s = 0 t \neq 0 \\ ax + \frac{(b-\delta)}{2} - h(sx) & \text{if } s \neq 0 t = 0 \\ 2ax^3 + bx^2 + cx - A(x) + \beta - \frac{\delta}{2} & \text{if } s = t \neq 0 \\ 3ax^2 + cx + \frac{1}{2}A_o(x) + \beta - d, & \text{if } s = -t \neq 0 \\ bt(t-2s)x^2 + \beta x + A(tx) + \gamma & \text{if } 0 \neq s^2 \neq t^2 \neq 0 \end{cases} \\
 h(x) &= \begin{cases} \text{arbitrary with } h(0) = d & \text{if } s = 0 = t \\ \text{arbitrary} & \text{if } s = 0, t \neq 0 \\ \text{arbitrary} & \text{if } s \neq 0, t = 0 \\ a\left(\frac{x}{t}\right)^3 + b\left(\frac{x}{t}\right)^2 + A\left(\frac{x}{t}\right) + d & \text{if } s = t \neq 0 \\ -a\left(\frac{x}{t}\right)^2 - \frac{t}{x}A\left(\frac{x}{t}\right) + \frac{1}{2}A_o\left(\frac{x}{t}\right), x \neq 0 & \text{if } s = -t \neq 0, \\ -bx^2 - A(x) - d & \text{if } 0 \neq s^2 \neq t^2 \neq 0, \end{cases}
 \end{aligned}$$

where $A_o, A : \mathbb{R} \rightarrow \mathbb{R}$ are additive functions and $a, b, c, d, \alpha, \beta, \gamma, \delta$ are arbitrary real constants.

Proof: Letting $x = y$ in (3.127), we see that

$$f(x) = g(x) \quad (3.188)$$

for all $x \in \mathbb{R}$. Hence (3.188) in (3.127) yields

$$f(x) - f(y) = (x - y)[h(sx + ty) + \phi(x) + \psi(y)]. \quad (3.189)$$

Interchanging x and y in (3.189) and adding the resulting equation to (3.189), we have

$$h(sx + ty) + \phi(x) + \psi(y) = h(sy + tx) + \phi(y) + \psi(x) \quad (3.190)$$

for all $x, y \in \mathbb{R}$ with $x \neq y$. But (3.190) holds even for $x = y$.

Now we consider several cases.

Case 1. Suppose $s = 0 = t$. Then (3.190) yields

$$\phi(x) = \psi(x) - \delta, \quad (3.191)$$

where δ is a constant. Letting (3.191) into (3.189), we have

$$f(x) - f(y) = (x - y)[h(sx + ty) + \psi(x) + \psi(y) - \delta]. \quad (3.192)$$

Hence by Theorem 3.10, (3.188) and (3.191) we obtain the asserted solution

$$\left. \begin{array}{l} f(x) = ax^2 + (b + d)x + c \\ g(x) = f(x) \\ \phi(x) = ax + \frac{b-\delta}{2} \\ \psi(x) = ax + \frac{b+\delta}{2} \\ h(x) \text{ arbitrary with } h(0) = d, \end{array} \right\}$$

where a, b, c, d, δ are arbitrary constants.

Case 2. Suppose $s = 0$ and $t \neq 0$. (The case $s \neq 0$ and $t = 0$ can be handled in a similar manner.) Then for this case, from (3.190), we have

$$h(ty) + \psi(y) - \phi(y) = h(tx) + \psi(x) - \phi(x) \quad (3.193)$$

for all $x, y \in \mathbb{R}$. Thus

$$\psi(x) = \phi(x) - h(tx) - \delta \quad (3.194)$$

where δ is a constant. Letting (3.194) in (3.189) with $s = 0$, we see that

$$f(x) - f(y) = (x - y)[\phi(x) + \phi(y) - \delta]. \quad (3.195)$$

By Theorem 3.8, (3.188) and (3.194), we have the asserted solution

$$\left. \begin{array}{l} f(x) = ax^2 + bx + c \\ g(x) = f(x) \\ \phi(x) = ax + \frac{b+\delta}{2} \\ \psi(x) = ax + \frac{b-\delta}{2} - h(tx) \\ h(x) \text{ arbitrary} \end{array} \right\}$$

where a, b, d, c, δ are arbitrary constants.

Case 3. Suppose $s \neq 0 \neq t$. Next, we consider several subcases.

Subcase 3.1. Suppose $s = t$. Then from (3.190), we get

$$h(tx + ty) + \phi(x) + \psi(y) = h(ty + tx) + \phi(y) + \psi(x). \quad (3.196)$$

Hence, we have

$$\phi(x) = \psi(x) - \delta, \quad (3.197)$$

where δ is a constant. Letting (3.197) into (3.189), we obtain

$$f(x) - f(y) = (x - y)[h(tx + ty) + \psi(x) + \psi(y) - \delta]. \quad (3.198)$$

From Theorem 3.10, (3.197) and (3.188), we obtain

$$\left. \begin{array}{l} f(x) = 3ax^4 + 2bx^3 + cx^2 + (d + 2\beta)x + \alpha \\ g(x) = f(x) \\ \phi(x) = 2ax^3 + bx^2 + cx - A(x) + \beta + \frac{\delta}{2} \\ \psi(x) = 2ax^3 + bx^2 + cx - A(x) + \beta - \frac{\delta}{2} \\ h(x) = a\left(\frac{x}{t}\right)^3 + b\left(\frac{x}{t}\right)^2 + A\left(\frac{x}{t}\right) + d, \end{array} \right\}$$

where $a, b, c, d, \alpha, \beta, \delta$ are arbitrary constants and $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function.

Subcase 3.2. Suppose $s = -t$. From (3.190), we have

$$h(ty - tx) + \phi(x) + \psi(y) = h(tx - ty) + \phi(y) + \psi(x) \quad (3.199)$$

for all $x, y \in \mathbb{R}$. This in turn yields

$$h(tx - ty) - h(ty - tx) = H(x) - H(y), \quad (3.200)$$

where $H(x) = \phi(x) - \psi(x)$. Letting $x = 0$ in (3.200), we observe that

$$h(-ty) - h(ty) = d - H(y), \quad (3.201)$$

where $d = H(0)$. Using (3.201) in (3.200), we have

$$H(x - y) + d = H(x) + d - H(y) - d,$$

that is $H(x) + d$ is additive on the set of reals. Hence

$$\psi(x) = \phi(x) + A_o(x) - d, \quad (3.202)$$

where $A_o : \mathbb{R} \rightarrow \mathbb{R}$ is an additive map. Substituting (3.202) into (3.189), we get

$$f(x) - f(y) = (x - y) [h(ty - tx) + \phi(x) + \phi(y) + A_o(y) - d] \quad (3.203)$$

which is

$$F(x) - F(y) = (x - y) [K(tx - ty) + \Phi(x) + \Phi(y)] \quad (3.204)$$

where

$$\left. \begin{aligned} F(x) &= f(x) + dx \\ K(x) &= h(-tx) - \frac{1}{2}A_o\left(\frac{x}{t}\right) \\ \Phi(x) &= \phi(x) + \frac{1}{2}A_o(x). \end{aligned} \right\} \quad (3.205)$$

Thus from Theorem 3.10, (3.188), (3.202) and (3.205), we again have the asserted solution

$$\left. \begin{aligned} f(x) &= 2ax^3 + cx^2 + (2\beta - d)x - A(x) + \alpha \\ g(x) &= f(x) \\ \phi(x) &= 3ax^2 + cx - \frac{1}{2}A_o(x) + \beta \\ \psi(x) &= 3ax^2 + cx + \frac{1}{2}A_o(x) + \beta - d \\ h(x) &= -a\left(\frac{x}{t}\right)^2 - \frac{t}{x}A\left(\frac{x}{t}\right) + \frac{1}{2}A_o\left(\frac{x}{t}\right), x \neq 0 \end{aligned} \right\}$$

where $a, b, c, d, \alpha, \beta$ are arbitrary constants and $A, A_o : \mathbb{R} \rightarrow \mathbb{R}$ are additive functions.

Subcase 3.3. Suppose $s^2 \neq t^2$. Letting $y = 0$ in (3.190), we get

$$h(sx) + \phi(x) + \psi(0) = h(tx) + \phi(0) + \psi(x). \quad (3.206)$$

Letting (3.206) in (3.190) and simplifying, we have

$$h(sx + ty) - h(sx) - h(ty) = h(sy + tx) - h(tx) - h(sy). \quad (3.207)$$

Replacing x by $\frac{x}{s}$ and y by $\frac{y}{t}$ in (3.189), we obtain

$$f\left(\frac{x}{s}\right) - f\left(\frac{y}{t}\right) = \left(\frac{xt - ys}{st}\right) [h(x + y) + \phi\left(\frac{x}{s}\right) + \psi\left(\frac{y}{t}\right)]. \quad (3.208)$$

Defining

$$\left. \begin{aligned} F(x) &= f\left(\frac{x}{s}\right) \\ \Phi(x) &= \phi\left(\frac{x}{s}\right) \\ \Psi(x) &= \psi\left(\frac{y}{t}\right) \end{aligned} \right\} \quad (3.209)$$

and using (3.209) in (3.208), we have

$$F(x) - F\left(\frac{sy}{t}\right) = (xt - ys)[h(x + y) + \Phi(x) + \Psi(y)]. \quad (3.210)$$

Letting $y = 0$ and $x = 0$ separately in (3.210), we have

$$F(x) = F(0) + xt[h(x) + \Phi(x) + \Psi(0)] \quad (3.211)$$

and

$$F\left(\frac{sy}{t}\right) = F(0) + ys[h(y) + \Phi(0) + \Psi(y)], \quad (3.212)$$

respectively. Letting (3.211) and (3.212) into (3.210), we obtain (after some simplifications)

$$\begin{aligned} & xt[\Psi(0) - \Psi(y) - h(y)] - ys[\Phi(0) - \Phi(x) - h(x)] \\ &= (xt - ys)[h(x + y) - h(x) - h(y)]. \end{aligned} \quad (3.213)$$

Interchanging x with y in (3.213), we obtain

$$\begin{aligned} & yt[\Psi(0) - \Psi(x) - h(x)] - xs[\Phi(0) - \Phi(y) - h(y)] \\ &= (yt - xs)[h(x + y) - h(x) - h(y)]. \end{aligned} \quad (3.214)$$

Subtracting (3.214) from (3.213), we have

$$\begin{aligned} & xP(y) - yP(x) \\ &= (x - y)(s + t)[h(x + y) - h(x) - h(y)], \end{aligned} \quad (3.215)$$

where

$$P(x) = t[\Phi(0) - \Phi(x) - h(x)] + s[\Psi(0) - \Psi(x) - h(x)]. \quad (3.216)$$

The general solution of (3.215) can be obtained from Theorem 3.7 as

$$\begin{array}{lcl} P(x) & = & 3ax^3 + 2bx^2 + cx + d \\ (s + t)h(x) & = & -ax^3 - bx^2 - A(x) - d, \end{array} \quad \left. \right\} \quad (3.217)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function and a, b, c, d are arbitrary constants. Letting the form of $h(x)$ in (3.217) in (3.207), we obtain

$$3a st xy(s - t)(x - y) = 0$$

for all $x, y \in \mathbb{R}$. Hence $a = 0$ as $0 \neq s^2 \neq t^2 \neq 0$. Thus, we have

$$(s+t)h(x) = -bx^2 - A(x) - d. \quad (3.218)$$

Using (3.218) in (3.206), we get

$$\phi(x) = \psi(x) + b(s-t)x^2 + \frac{A(sx-tx)}{s+t} + \alpha, \quad (3.219)$$

where $\alpha = \phi(0) - \psi(0)$. Letting (3.218) and (3.219) into (3.189), we obtain

$$k(x) - k(y) = (x-y)[\alpha_0xy + \Gamma(x) + \Gamma(y)], \quad (3.220)$$

where

$$\left. \begin{aligned} k(x) &= f(x) + \left(\frac{d}{s+t} - \alpha \right) x \\ \Gamma(x) &= \psi(x) - \frac{A(tx)}{s+t} - \frac{bt^2x^2}{s+t} \\ \alpha_0 &= -\frac{2bst}{s+t}. \end{aligned} \right\} \quad (3.221)$$

Using Lemma 3.3, we obtain

$$k(x) = \alpha_0x^3 + \beta x^2 + 2\gamma x + \delta \quad \text{and} \quad \Gamma(x) = \alpha_0x^2 + \beta x + \gamma \quad (3.222)$$

where β, γ, δ are arbitrary constants. Hence from (3.222), (3.221), (3.219) and (3.218), we have

$$\begin{aligned} f(x) &= -\frac{2bstx^3}{s+t} + \beta x^2 + \left(2\gamma + \alpha - \frac{d}{s+t} \right) x + \delta \\ \phi(x) &= \frac{bs(s-2t)x^2}{s+t} + \beta x + \frac{A(sx)}{s+t} + \gamma + \alpha \\ \psi(x) &= \frac{bt(t-2s)x^2}{s+t} + \beta x + \frac{A(tx)}{s+t} + \gamma \\ h(x) &= -\frac{bx^2}{s+t} - \frac{A(x)}{s+t} - \frac{d}{s+t}. \end{aligned}$$

Renaming the constants $\frac{b}{s+t}$ as b , $\frac{d}{s+t}$ as d , and the additive function $\frac{A(x)}{s+t}$ as $A(x)$, we have the asserted solution

$$\begin{aligned} f(x) &= -2bstx^3 + \beta x^2 + (2\gamma + \alpha - d)x + \delta \\ \phi(x) &= bs(s-2t)x^2 + \beta x + A(sx) + \gamma + \alpha \\ \psi(x) &= bt(t-2s)x^2 + \beta x + A(tx) + \gamma \\ h(x) &= -bx^2 - A(x) - d. \end{aligned}$$

Since no more cases are left, the proof of the theorem is now complete.

3.6 Some Open Problems

The following problem is due to Kuczma (1991). Let $[a, b]$ be a proper interval, that is an interval of finite length, such that $1 \in]a, b[$. Find the general solution of the functional equation

$$\begin{aligned} \psi(y) [1 - \psi(x)] \phi\left(\frac{1+x}{2}\right) - \psi(x) [1 - \psi(y)] \phi\left(\frac{1+y}{2}\right) \\ = [\psi(y) - \psi(x)] \phi\left(\frac{x+y}{2}\right) \end{aligned}$$

for all $x, y \in [a, b]$. In particular, it will be interesting to know if this functional equation has a continuous solution $\phi : [a, b] \rightarrow \mathbb{R}$ which is neither constant nor strictly monotonic on the open interval $]a, b[$.

In section 3.4, we determined the general solution of the functional equation

$$f(x) - g(y) = [x - y] [h(x + y) + \psi(x) + \phi(y)]$$

for all $x, y \in \mathbb{R}$. The factor $x - y$ can be replaced by $k(x) - k(y)$. Of course, we would like the function $k : \mathbb{R} \rightarrow \mathbb{R}$ to be continuous and strictly increasing. Thus, our second problem is the following: Find all real valued functions $f, g, h, k, \psi, \phi : \mathbb{R} \rightarrow \mathbb{R}$ satisfying the functional equation

$$f(x) - g(y) = [k(x) - k(y)] [h(x + y) + \psi(x) + \phi(y)], \quad \forall x, y \in \mathbb{R}.$$

In particular, we are interested in the solution of the above functional equation when the function $k : \mathbb{R} \rightarrow \mathbb{R}$ is continuous and strictly increasing.

Note that the functional equation

$$f(x) - f(y) = (x - y) [h(sx + ty) + \phi(x) + \psi(y)] \quad (3.223)$$

can be generalized to

$$f[x_1, x_2, \dots, x_n] = h\left(\sum_{i=1}^n w_i x_i\right) + \sum_{i=1}^n \phi_i(x_i), \quad (3.224)$$

where x_1, x_2, \dots, x_n are distinct real numbers and w_1, w_2, \dots, w_n are given real parameters. Find all real valued functions $f, h, \phi_1, \dots, \phi_n$ on the real line which satisfy the functional equation (3.224) for distinct $x_1, x_2, \dots, x_n \in \mathbb{R}$.

The functional equation $f\{x, y\} = h(x + y)$ can be generalized to

$$f\{x_1, x_2, \dots, x_n\} = h\left(g^{-1}\left(\frac{1}{n} \sum_{i=1}^n g(x_i)\right)\right), \quad (3.225)$$

where g is a continuous and strictly monotone function. Determine the general solution of the functional equation (3.225).

Similar to equation (3.223), using the trapezoidal rule for numerical integration, we have the functional equation

$$f(x) - f(y) = (x - y)[g(x) + h(x + 2y) + h(2x + y) + g(y)] \quad (3.226)$$

for all $x, y \in \mathbb{R}$. Determine the general solution of equation (3.226).

Chapter 4

Two-dimensional Mean Value Theorems and Functional Equations

Lagrange's mean value theorem for functions in one variable can be extended to functions in two variables or more. The main goal of this chapter is to present some results regarding the mean value theorem for functions in two variables and then to discuss Cauchy's mean value theorem for functions in two variables. Further, we study some functional equations arising from the mean value theorem for functions in two variables. The functional equations studied in this chapter are related to the characterizations of low degree polynomials in two variables. Finally, we point out some open problems in the area of functional equations that are related to our topic of discussion.

4.1 MVTs for Functions in Two Variables

The following result appeared in the book by Courant (1964).

Theorem 4.1 *For every function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ with continuous partial derivatives f_x and f_y and for all distinct pairs (x, y) and (u, v) in \mathbb{R}^2 , there exists an intermediate point (η, ξ) on the line segment joining the points (x, y) and (u, v) such that*

$$f(u, v) - f(x, y) = (u - x)f_x(\eta, \xi) + (v - y)f_y(\eta, \xi). \quad (4.1)$$

Proof: Let (x, y) and (u, v) be any two points in the plane \mathbb{R}^2 . Let $h = u - x$ and $k = v - y$. Let \mathcal{L} be line-segment obtained by joining the points (x, y) and (u, v) . The co-ordinates of any point on this line-segment are given by

$(x + ht, y + kt)$ for some $t \in [0, 1]$. We define a function $F : [0, 1] \rightarrow \mathbb{R}$ by

$$F(t) = f(x + ht, y + kt)$$

keeping x, y, u, v fixed for the moment. The derivative of this function is given by

$$F'(t) = h f_x(x + ht, y + kt) + k f_y(x + ht, y + kt), \quad (4.2)$$

where f_x and f_y are the partial derivatives of f with respect to x and y , respectively. Applying the mean value theorem to F yields,

$$F(1) - F(0) = F'(t_o), \quad (4.3)$$

where $t_o \in]0, 1[$. The definition of F in (4.3) yields

$$f(u, v) - f(x, y) = F'(t_o).$$

Using (4.2) in the above equation, we get

$$f(u, v) - f(x, y) = h f_x(x + ht_o, y + kt_o) + k f_y(x + ht_o, y + kt_o)$$

which is

$$f(u, v) - f(x, y) = (u - x) f_x(\eta, \xi) + (v - y) f_y(\eta, \xi),$$

where (η, ξ) is the point on line-segment \mathcal{L} whose co-ordinates are given by $(x + ht_o, y + kt_o)$. This completes the proof of the theorem.

The geometrical interpretation of this theorem is that the difference between the values of the function at the points (u, v) and (x, y) is equal to the differential at an intermediate point (η, ξ) on the line-segment joining the two points.

This mean value theorem can be used to deduce the following fact. A function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ whose partial derivatives f_x and f_y exist and have the value 0 at every point on the plane is constant. To see this, let (x, y) and (u, v) be any two arbitrary points on the plane and apply the above theorem to f . Then we have

$$f(u, v) - f(x, y) = (u - x) f_x(\eta, \xi) + (v - y) f_y(\eta, \xi),$$

for some (η, ξ) on the line-segment joining the points (x, y) and (u, v) . Since the partial derivatives are zero at every points, we see that

$$f(x, y) = f(u, v)$$

for all $x, y, u, v \in \mathbb{R}$. Therefore f is a constant function on \mathbb{R}^2 .

4.2 Mean Value Type Functional Equations

The equation (4.1) yields a functional equation of the form

$$f(u, v) - f(x, y) = (u - x)g(\eta, \xi) + (v - y)h(\eta, \xi), \quad (4.4)$$

for all $x, y, u, v \in \mathbb{R}$ with $(u - x)^2 + (v - y)^2 \neq 0$. In (4.4), f_x and f_y are replaced by unknown functions g and h , respectively.

If $\eta(u, x) = u + x$ and $\xi(v, y) = v + y$, then (4.1) reduces to

$$\begin{aligned} & f(u, v) - f(x, y) \\ &= (u - x)g(u + x, v + y) + (v - y)h(u + x, v + y), \end{aligned} \quad (4.5)$$

for all $x, y, u, v \in \mathbb{R}$ with $(u - x)^2 + (v - y)^2 \neq 0$. This equation was investigated in Kannappan and Sahoo (1998) to characterize quadratic functions in two variables. We now prove some lemmas to establish the next theorem.

Lemma 4.1 *The functions $f, g, \ell : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the equation*

$$f(x) - g(y) = (x - y)\ell(x + y) \quad (4.6)$$

for all $x, y \in \mathbb{R}$ if, and only if, there exist constants a, b , and c such that

$$f(x) = ax^2 + bx + c, \quad g(x) = ax^2 + bx + c, \quad \ell(x) = ax + b. \quad (4.7)$$

Proof: Interchanging x with y in (4.6) and using (4.6), we have

$$f(x) - g(x) = g(y) - f(y) \quad (4.8)$$

for all $x, y \in \mathbb{R}$. Hence $f(x) = g(x) + c$, where c is a constant. Letting this into (4.8), we have $c = -c$ and hence $c = 0$. Therefore $f(x) = g(x)$. Now using Theorem 2.3, we have the asserted solution (4.7) and the proof is complete.

Lemma 4.2 *The general solution $f, g, \ell : \mathbb{R} \rightarrow \mathbb{R}$ of the generalized Jensen functional equation*

$$f(x + y) + g(x - y) = \ell(x) \quad (4.9)$$

is given by

$$f(x) = A(x) + b, \quad g(x) = A(x) + c, \quad \ell(x) = 2A(x) + b + c, \quad (4.10)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function, and b and c are arbitrary constants.

Proof: Letting $y = 0$ in (4.9), we have

$$\ell(x) = f(x) + g(x). \quad (4.11)$$

In view of (4.11), (4.9) becomes

$$f(x+y) + g(x-y) = f(x) + g(x) \quad (4.12)$$

Replacing y by $-y$ in (4.12) and then using (4.12), we see that

$$f(x+y) - g(x+y) = f(x-y) - g(x-y)$$

for all $x, y \in \mathbb{R}$. Hence

$$f(x) = g(x) + c_0 \quad (4.13)$$

where c_0 is an arbitrary constant. Letting (4.13) into (4.12) and defining

$$F(x) = f(x) - f(0), \quad (4.14)$$

we have

$$F(x+y) + F(x-y) = 2F(x) \quad (4.15)$$

with $F(0) = 0$. Letting $x = 0$ in (4.15), we conclude that F is an odd function in \mathbb{R} . Interchanging x with y in (4.15) and using the fact that F is an odd function, we get

$$F(x+y) - F(x-y) = 2F(y). \quad (4.16)$$

Adding (4.15) and (4.16), we get

$$F(x+y) = F(x) + F(y), \quad x, y \in \mathbb{R},$$

that is $F(x) = A(x)$, where $A : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function. Now using (4.14) and (4.11), we have the asserted solution (4.10). This completes the proof.

The functional equation (4.9) in the above lemma is a generalization of the Jensen functional equation

$$f\left(\frac{u+v}{2}\right) = \frac{f(u)+f(v)}{2}$$

for all $u, v \in \mathbb{R}$. To see this let $\ell = 2f$ and $g = f$, and replace $x + y$ by u and $x - y$ by v in (4.9) to obtain the Jensen functional equation. The continuous solution of the Jensen equation is affine linear, that is, of the form $f(x) = mx + k$, where m, k are arbitrary real constants.

Lemma 4.3 *The function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ satisfies the functional equation*

$$\begin{aligned} & f(u+x, v+y) + f(u-x, v) + f(u, v-y) \\ &= f(u-x, v-y) + f(u+x, v) + f(u, v+y) \end{aligned} \quad (4.17)$$

for all $x, y, u, v \in \mathbb{R}$ if, and only if,

$$f(x, y) = B(x, y) + \phi(x) + \psi(y), \quad (4.18)$$

where $B : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a biadditive function, and $\phi, \psi : \mathbb{R} \rightarrow \mathbb{R}$ are arbitrary functions.

Proof: For fixed v and y , we define

$$\left. \begin{array}{lcl} F(x) & = & f(x, v+y) \\ G(x) & = & f(x, v), \\ H(x) & = & f(x, v-y). \end{array} \right\} \quad (4.19)$$

Then by (4.19), we see that (4.17) reduces to

$$(F - G)(u+x) + (G - H)(u-x) = (F - H)(u). \quad (4.20)$$

By Lemma 4.2, the general solution of (4.20) is given by

$$\left. \begin{array}{lcl} (F - G)(u) & = & A(u) + \beta \\ (G - H)(u) & = & A(u) + \alpha \\ (F - H)(u) & = & 2A(u) + \alpha + \beta, \end{array} \right\} \quad (4.21)$$

where A is an additive function, and α and β are arbitrary real constants all dependent on v and y . From (4.19) and (4.21), we obtain

$$f(x, y) - f(x, 0) = A(x, y) + \beta(y)$$

which is

$$f(x, y) = A(x, y) + \alpha(x) + \beta(y), \quad (4.22)$$

where $A : \mathbb{R}^2 \rightarrow \mathbb{R}$ is additive in the first variable, and $\alpha(x) = f(x, 0)$. Substituting (4.22) into (4.17) and simplifying, we obtain

$$A(x, v+y) + A(x, v-y) = 2A(x, v) \quad (4.23)$$

for all $x, y, v \in \mathbb{R}$. For a fixed x , (4.23) is a Jensen equation and using Lemma 4.2 we have $A(x, y) = B(x, y) + C(x)$, where B is additive in the second variable y . Since B is additive in the second variable, we have

$$A(x, y_1 + y_2) = A(x, y_1) + A(x, y_2) - C(x).$$

Using the additivity of A in the first variable, we see that C is additive, and hence B is additive in first variable also. Thus B is biadditive and f is of the form 4.18 as asserted. This establishes the lemma.

Theorem 4.2 *Let $f, g_1, g_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ satisfy*

$$\begin{aligned} & f(u, v) - f(x, y) \\ &= (u-x)g_1(x+u, y+v) + (v-y)g_2(x+u, y+v), \end{aligned} \quad (4.24)$$

for all $x, y, u, v \in \mathbb{R}$ with $(u-x)^2 + (v-y)^2 \neq 0$, then f is of the form

$$f(x, y) = B(x, y) + ax^2 + bx + cy^2 + dy + \alpha, \quad (4.25)$$

where $B : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a biadditive function, and a, b, c, d, α are arbitrary constants.

Proof: Letting $u = -x$ and $v = -y$ in (4.24), we get

$$f(-x, -y) = f(x, y) - a_1x - b_1y \quad (4.26)$$

where $a_1 = 2g_1(0, 0)$ and $b_1 = 2g_2(0, 0)$. Replacing x by $-x$ and y by $-y$ in (4.24), we get

$$\begin{aligned} & f(u, v) - f(-x, -y) \\ &= (u+x)g_1(u-x, v-y) + (v+y)g_2(u-x, v-y). \end{aligned} \quad (4.27)$$

Using (4.26) and (4.27), we obtain

$$\begin{aligned} f(u, v) - f(x, y) + a_1x + b_1y \\ = (u + x)g_1(u - x, v - y) + (v + y)g_2(u - x, v - y). \end{aligned} \quad (4.28)$$

Comparing (4.24) with (4.28), we see that

$$\begin{aligned} (u + x)g_1(u - x, v - y) + (v + y)g_2(u - x, v - y) \\ = (u - x)g_1(u + x, v + y) + (v - y)g_2(u + x, v + y) + a_1x + b_1y. \end{aligned}$$

Substituting $u = 1 + x$ and $v = 1 + y$ in the above equation, we get

$$\begin{aligned} (2x + 1)g_1(1, 1) + (2y + 1)g_2(1, 1) \\ = g_1(2x + 1, 2y + 1) + g_2(2x + 1, 2y + 1) + a_1x + b_1y. \end{aligned} \quad (4.29)$$

From (4.29), we note that

$$g_2(u, v) = a_2u + b_2v + c_2 - g_1(u, v), \quad (4.30)$$

where a_2, b_2, c_2 are real constants. Replacing u by $u + x$, v by $v + y$, x by $u - x$, and y by $v - y$ in (4.24), we obtain

$$\begin{aligned} f(u + x, v + y) - f(u - x, v - y) \\ = 2xg_1(2u, 2v) + 2yg_2(2u, 2v). \end{aligned} \quad (4.31)$$

By (4.30), equation (4.31) yields

$$\begin{aligned} f(u + x, v + y) - f(u - x, v - y) \\ = 2(x - y)g_1(2u, 2v) + 4a_2uy + 4b_2vy + 2c_2y. \end{aligned} \quad (4.32)$$

Now, putting $y = 0$ in (4.32), we see that

$$f(u + x, v) - f(u - x, v) = 2xg_1(2u, 2v). \quad (4.33)$$

Further, $x = 0$ in (4.32) yields

$$\begin{aligned} f(u, v + y) - f(u, v - y) \\ = -2yg_1(2u, 2v) + 4a_2uy + 4b_2vy + 2c_2y. \end{aligned} \quad (4.34)$$

Adding (4.33) to (4.34) and subtracting the resulting equation from (4.32), we get

$$\begin{aligned} f(u+x, v+y) + f(u-x, v) + f(u, v-y) \\ = f(u-x, v-y) + f(u+x, v) + f(u, v+y). \end{aligned} \quad (4.35)$$

The general solution of (4.35) can be obtained from Lemma 4.3 as

$$f(x, y) = B(x, y) + \phi(x) + \psi(y), \quad (4.36)$$

where $B : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a biadditive function, and $\phi, \psi : \mathbb{R} \rightarrow \mathbb{R}$ are arbitrary functions. Next, we let $v = y = 0$ in (4.24) to get

$$f(u, 0) - f(x, 0) = (u-x)g_1(u+x, 0), \quad u \neq x. \quad (4.37)$$

Now using Lemma 4.1, we obtain

$$f(x, 0) = ax^2 + bx + \alpha_1, \quad (4.38)$$

where a, b, α_1 are arbitrary constants. Letting $y = 0$ in (4.36) and comparing with (4.38), we get

$$\phi(x) = ax^2 + bx + \alpha_2, \quad (4.39)$$

where α_2 is a constant. Again, letting $u = x = 0$ in (4.24), we get

$$f(0, v) - f(0, y) = (v-y)g_2(0, v+y), \quad v \neq y. \quad (4.40)$$

Again, the use of Lemma 4.1 gives

$$f(0, y) = cy^2 + dy + \alpha_3, \quad (4.41)$$

where c, d, α_3 are arbitrary constants. As before, letting $x = 0$ in (4.36) and comparing with (4.41), we get

$$\psi(y) = cy^2 + dy + \alpha_4, \quad (4.42)$$

where α_4 is a constant. Now (4.39), (4.42) and (4.36) yield the asserted solution (4.25), that is

$$f(x, y) = B(x, y) + ax^2 + bx + cy^2 + dy + \alpha,$$

where $\alpha = \alpha_2 + \alpha_4$. This completes the proof of Theorem 4.2.

Using the above theorem Kannappan and Sahoo (1998) established the following property of quadratic functions in two variables analogous to Theorem 2.5 of Chapter 2.

Theorem 4.3 *If the quadratic polynomial $f(x, y) = ax^2 + bx + cy^2 + dy + exy + \alpha$ with $4ac - e^2 \neq 0$, is a solution of the functional differential equation*

$$\begin{aligned} f(x + h, y + k) - f(x, y) \\ = h f_x(x + \theta h, y + \theta k) + k f_y(x + \theta h, y + \theta k), \end{aligned} \quad (4.43)$$

assumed for all $x, y, h, k \in \mathbb{R}$ with $h^2 + k^2 \neq 0$, then $\theta = \frac{1}{2}$. Conversely, if a function f satisfies (4.43) with $\theta = \frac{1}{2}$, then the only solution is at most a quadratic polynomial.

Proof: Suppose the polynomial $f(x, y) = ax^2 + bx + cy^2 + dy + exy + \alpha$ is a solution of (4.43). Then from (4.43), we obtain

$$2ah^2\theta + 2ck^2\theta + 2ehk\theta = ah^2 + ck^2 + ehk.$$

Since $4ac - e^2 \neq 0$, we obtain

$$\theta = \frac{1}{2}.$$

This proves the ‘if’ part of the theorem.

Next, we prove the converse. By Theorem 4.2, we get

$$f(x, y) = B(x, y) + ax^2 + bx + cy^2 + dy + \alpha. \quad (4.44)$$

Since f has partial derivatives with respect to x and y , so also B . Thus

$$B(x, y) = exy, \quad (4.45)$$

since B is biadditive. Here e is an arbitrary constant. Letting (4.45) in (4.44), we see that f is a quadratic polynomial. This completes the proof of Theorem 4.3.

4.3 Generalized Mean Value Type Equations

Recently, Riedel and Sahoo (1997) found the general solution of the functional equation (4.4) when $\eta = sx + tu$ and $\xi = sy + tv$ for real *a priori*

chosen parameters s and t . We summarize their result in the following theorem.

Theorem 4.4 *The functions f, g , and h from \mathbb{R}^2 to \mathbb{R} are solutions of the functional equation*

$$\begin{aligned} & f(u, v) - f(x, y) \\ &= (u - x)g(sx + tu, sy + tv) + (v - y)h(sx + tu, sy + tv), \end{aligned} \quad (4.46)$$

for all x, u, v, y in \mathbb{R} with $(u - x)^2 + (v - y)^2 \neq 0$, if and only if

$$\left. \begin{array}{l} f(x, y) = ax + by + c \\ g(x, y) = \text{arbitrary with } g(0, 0) = a \\ h(x, y) = \text{arbitrary with } h(0, 0) = b \end{array} \right\} \text{if } s = t = 0, \quad (4.47)$$

$$\left. \begin{array}{l} f(x, y) = ax + by + c \\ g(x, y) = a \\ h(x, y) = b \end{array} \right\} \text{if } s = 0 \text{ or } t = 0, \text{ but not both, or } s^2 \neq t^2 \quad (4.48)$$

$$\left. \begin{array}{l} f(x, y) = (\alpha - \gamma)x^2t + 2\gamma xy t \\ \quad + \delta y^2 t + ax + by + c \\ g(x, y) = (\alpha - \gamma)x + \gamma y + a \\ h(x, y) = \gamma x + \delta y + b \end{array} \right\} \text{if } s = t \neq 0, \quad (4.49)$$

$$\left. \begin{array}{ll} f(x, y) &= \frac{A(tx)}{t} + \frac{B(ty)}{t} + c \\ g(x, y) &= \frac{A(x)}{x} + \frac{B(y)}{x} - \frac{y}{x}h(x, y) \\ h(x, y) &= \text{arbitrary} \\ \text{for } x \neq 0; \text{ or for } y \neq 0 & \\ f(x, y) &= \frac{A(tx)}{t} + \frac{B(ty)}{t} + c \\ g(x, y) &= \text{arbitrary} \\ h(x, y) &= \frac{A(y)}{y} + \frac{B(x)}{y} - \frac{x}{y}g(x, y) \end{array} \right\} \text{if } s = -t \neq 0, \quad (4.50)$$

where $A : \mathbb{R} \rightarrow \mathbb{R}$ and $B : \mathbb{R} \rightarrow \mathbb{R}$ are additive functions, and $a, b, c, \alpha, \delta, \gamma$ are arbitrary, real constant.

Proof. We prove this theorem by considering the various cases of the parameters s and t .

Case 1: Suppose $s = t = 0$. Then (4.46) reduces to

$$f(u, v) - f(x, y) = (u - x)a + (v - y)b, \quad (4.51)$$

where $a = g(0, 0)$ and $b = h(0, 0)$. From (4.51) we obtain

$$f(u, v) - au - bv = f(x, y) - ax - by \quad (4.52)$$

for all $x, y, u, v \in \mathbb{R}$ with $(u - x)^2 + (v - y)^2 \neq 0$. Hence,

$$f(x, y) = ax + by + c \quad (4.53)$$

where c is an arbitrary real constant. Thus the solution of (4.46) becomes

$$\left. \begin{array}{l} f(x, y) = ax + by + c \\ g(x, y) = \text{arbitrary with } g(0, 0) = a \\ h(x, y) = \text{arbitrary with } h(0, 0) = b, \end{array} \right\} \quad (4.54)$$

where a, b, c are arbitrary real constants.

Case 2: Suppose $t = 0$ or $s = 0$ but not both. Without loss of generality we assume $t = 0$. Then (4.46) yields

$$f(u, v) - f(x, y) = (u - x)g(sx, sy) + (v - y)h(sx, sy). \quad (4.55)$$

Letting $x = 0 = y$ in (4.55), we get

$$f(u, v) = au + bv + c, \quad (4.56)$$

where $a = g(0, 0)$, $b = h(0, 0)$ and $c = f(0, 0)$. Substituting this into (4.46), we obtain

$$a(u - x) + b(v - y) = (u - x)g(sx, sy) + (v - y)h(sx, sy), \quad (4.57)$$

for all $x, y, u, v \in \mathbb{R}$ with $(u - x)^2 + (v - y)^2 \neq 0$. Rewriting (4.57), we get

$$\begin{aligned} & u[a - g(sx, sy)] + v[b - h(sx, sy)] \\ & - \{x[a - g(sx, sy)] + y[b - h(sx, sy)]\} = 0. \end{aligned}$$

Using the linear independence of $u, v, 1$, we observe that

$$g(sx, sy) = a \quad (4.58)$$

and

$$h(sx, sy) = b, \quad (4.59)$$

for all $x, y \in \mathbb{R}$. Hence

$$\begin{aligned} g(x, y) &= a \\ h(x, y) &= b. \end{aligned} \quad \left. \right\} \quad (4.60)$$

And thus the solution of (4.46) becomes

$$\begin{aligned} f(x, y) &= ax + by + c \\ g(x, y) &= a \\ h(x, y) &= b, \end{aligned} \quad \left. \right\} \quad (4.61)$$

where a, b, c are arbitrary real constants.

Case 3: Suppose $s \neq 0 \neq t$. Substituting $x = 0 = y$ in (4.46), we obtain

$$f(u, v) = ug(tu, tv) + vh(tu, tv) + c, \quad (4.62)$$

where $c = f(0, 0)$. Letting (4.62) into (4.46), we see that

$$\begin{aligned} ug(tu, tv) + vh(tu, tv) - xg(tx, ty) - yh(tx, ty) \\ = (u - x)g(sx + tu, sy + tv) + (v - y)h(sx + tu, sy + tv). \end{aligned} \quad (4.63)$$

Replacing u by $\frac{u}{t}$, v by $\frac{v}{t}$, x by $\frac{x}{s}$ and y by $\frac{y}{s}$, in (4.63), we get

$$\begin{aligned} \frac{u}{t}g(u, v) + \frac{v}{t}h(u, v) - \frac{x}{s}g\left(\frac{tx}{s}, \frac{ty}{s}\right) - \frac{y}{s}h\left(\frac{xt}{s}, \frac{ty}{s}\right) \\ = \left(\frac{u}{t} - \frac{x}{s}\right)g(x + u, y + v) + \left(\frac{v}{t} - \frac{y}{s}\right)h(x + u, y + v), \end{aligned} \quad (4.64)$$

for all x, u, v, y in \mathbb{R} with $(us - xt)^2 + (vs - yt)^2 \neq 0$. Now consider several subcases:

Subcase 3.1: Suppose $s = t \neq 0$. Then (4.64) yields

$$\begin{aligned} ug(u, v) + vh(u, v) - xg(x, y) - yh(x, y) \\ = (u - x)g(u + x, v + y) + (v - y)h(u + x, v + y). \end{aligned} \quad (4.65)$$

Replacing x with $-x$ and y with $-y$ in (4.65), we obtain

$$\begin{aligned} ug(u, v) + vh(u, v) + xg(-x, -y) + yh(-x, -y) \\ = (u+x)g(u-x, v-y) + (v+y)h(u-x, v-y). \end{aligned} \quad (4.66)$$

Subtracting (4.65) from (4.66), we obtain

$$\begin{aligned} xg(-x, -y) + yh(-x, -y) + xg(x, y) + yh(x, y) \\ = (u+x)g(u-x, v-y) + (v+y)h(u-x, v-y) \\ - (u-x)g(u+x, v+y) - (v-y)h(u+x, v+y). \end{aligned} \quad (4.67)$$

Letting $u = -x$ and $v = -y$ in (4.65) and simplifying the result we get

$$xg(-x, -y) + yh(-x, -y) + xg(x, y) + yh(x, y) = 2ax + 2by, \quad (4.68)$$

where $a = g(0, 0)$ and $b = h(0, 0)$. Now we use (4.68) in (4.67) to obtain

$$\begin{aligned} 2ax + 2by \\ = (u+x)g(u-x, v-y) + (v+y)h(u-x, v-y) \\ - (u-x)g(u+x, v+y) - (v-y)h(u+x, v+y), \end{aligned} \quad (4.69)$$

which is

$$\begin{aligned} a(x+u) - a(u-x) + b(v+y) - b(v-y) \\ = (u+x)g(u-x, v-y) + (v+y)h(u-x, v-y) \\ - (u-x)g(u+x, v+y) - (v-y)h(u+x, v+y). \end{aligned}$$

Hence, we have

$$\begin{aligned} (u+x)[g(u-x, v-y) - a] + (v+y)[h(u-x, v-y) - b] \\ = (u-x)[g(u+x, v+y) - a] + (v-y)[h(u+x, v+y) - b]. \end{aligned} \quad (4.70)$$

Letting $u+x = 1 = v+y$ in (4.70), we get

$$g_0(u-x, v-y) + h_0(u-x, v-y) = \alpha(u-x) + \beta(v-y), \quad (4.71)$$

where $g_0 = g - a$ and $h_0 = h - b$. Letting (4.71) into (4.70), we see that

$$\begin{aligned} [(v+y) - (u+x)]h_0(u-x, v-y) + \beta(u+x)(v-y) \\ = [(v-y) - (u-x)]h_0(u+x, v+y) + \beta(u-x)(v+y), \end{aligned} \quad (4.72)$$

which can be written as

$$\begin{aligned} & [(v+y) - (u+x)][h_0(u-x, v-y) - \beta(v-y)] \\ & = [(v-y) - (u-x)][h_0(u+x, v+y) - \beta(v+y)]. \end{aligned} \quad (4.73)$$

Fixing $v+y$ and $u+x$ such that $v+y \neq u+x$ and separating variables gives

$$h_0(u-x, v-y) - \beta(v-y) = \alpha_0[(v-y) - (u-x)], \quad (4.74)$$

where α_0 is a constant. Thus we have that

$$h(u-x, v-y) = (\alpha_0 + \beta)(v-y) - \alpha_0(u-x) + b,$$

which is

$$h(x, y) = \delta y + \gamma x + b, \quad (4.75)$$

where γ, δ are constants. Letting (4.75) in (4.71), we get

$$g(x, y) = (\alpha - \gamma)x + (\beta - \delta)y + a. \quad (4.76)$$

Using (4.75) and (4.76) in (4.62), we get

$$f(x, y) = (\alpha - \gamma)x^2t + (\beta - \delta + \gamma)xyt + \delta y^2t + ax + by + c.$$

Thus we have the solution

$$\left. \begin{array}{l} f(x, y) = (\alpha - \gamma)x^2t + (\beta - \delta + \gamma)xyt \\ \quad + \delta y^2t + ax + by + c \\ g(x, y) = (\alpha - \gamma)x + (\beta - \delta)y + a \\ h(x, y) = \delta y + \gamma x + b. \end{array} \right\} \quad (4.77)$$

Substituting the above into (4.46) with $s = t \neq 0$ we obtain that

$$\beta - \delta = \gamma \quad (4.78)$$

which changes (4.77) to

$$\left. \begin{array}{l} f(x, y) = (\alpha - \gamma)x^2t + 2\gamma xy + \delta y^2t + ax + by + c \\ g(x, y) = (\alpha - \gamma)x + \gamma y + a \\ h(x, y) = \gamma x + \delta y + b. \end{array} \right\} \quad (4.79)$$

Subcase 3.2: Suppose $s = -t \neq 0$. Then (4.64) reduces to

$$\begin{aligned} ug(u, v) + vh(u, v) + xg(-x, -y) + yh(-x, -y) \\ = (u + x)g(u + x, v + y) + (v + y)h(u + x, v + y). \end{aligned} \quad (4.80)$$

Letting $u = 0 = v$ in (4.80), we obtain

$$xg(-x, -y) + yh(-x, -y) = xg(x, y) + yh(x, y). \quad (4.81)$$

Hence, using (4.81) in (4.80), we get

$$\begin{aligned} ug(u, v) + vh(u, v) + xg(x, y) + yh(x, y) = \\ (u + x)g(u + x, v + y) + (v + y)h(u + x, v + y). \end{aligned}$$

Letting $y = v = 0$ in (4.3), we see that

$$ug(u, 0) + xg(x, 0) = (u + x)g(u + x, 0).$$

Hence

$$ug(u, 0) = A(u), \quad (4.82)$$

where A is an arbitrary additive function. Similarly, letting $x = u = 0$ in (4.3), we get

$$vh(0, v) + yh(0, y) = (v + y)h(0, v + y).$$

Therefore

$$vh(0, v) = B(v), \quad (4.83)$$

where B is an arbitrary additive function. Next, letting $x = 0 = v$ in (4.3), we obtain

$$ug(u, 0) + yh(0, y) = ug(u, y) + yh(u, y). \quad (4.84)$$

Using (4.82) and (4.83) in (4.84), we have

$$ug(u, y) + yh(u, y) = A(u) + B(y). \quad (4.85)$$

Using (4.85) in (4.62), we obtain for $x \neq 0$,

$$\left. \begin{aligned} f(x, y) &= \frac{A(tx)}{t} + \frac{B(ty)}{t} + c \\ g(x, y) &= \frac{A(x)}{x} + \frac{B(y)}{x} - \frac{y}{x}h(x, y) \\ h(x, y) &= \text{arbitrary}, \end{aligned} \right\} \quad (4.86)$$

Subcase 3.3: Suppose $0 \neq s^2 \neq t^2 \neq 0$. We interchange x with u and y with v in (4.63) to get

$$\begin{aligned} & xg(tx, ty) + yh(tx, ty) - ug(tu, tv) - vh(tu, tv) = \\ & (x-u)g(tx+su, ty+sv) + (y-v)h(tx+su, ty+sv). \end{aligned}$$

Adding (4.63) to (4.3), we get

$$\begin{aligned} & (x-u)g(sx+tu, sy+tv) + (y-v)h(sx+tu, sy+tv) \\ & = (x-u)g(tx+su, ty+sv) + (y-v)h(tx+su, ty+sv). \quad (4.87) \end{aligned}$$

Letting $tu + sx = 0 = sy + tv$ in (4.87), we obtain

$$xg\left(\frac{t^2-s^2}{t}x, \frac{t^2-s^2}{t}y\right) + yh\left(\frac{t^2-s^2}{t}x, \frac{t^2-s^2}{t}y\right) = ax + by, \quad (4.88)$$

where $a = g(0, 0)$ and $b = h(0, 0)$. Replacing x by $\frac{tx}{t^2-s^2}$ and y by $\frac{ty}{t^2-s^2}$ in (4.88), we get

$$xg(x, y) + yh(x, y) = ax + by. \quad (4.89)$$

Using (4.89) in (4.62), we get

$$f(x, y) = ax + by + c. \quad (4.90)$$

Substituting (4.90) into (4.46) yields

$$\begin{aligned} & au + bv - ax - by \\ & = (u-x)g(sx+tu, sy+tv) + (v-y)h(sx+tu, sy+tv). \end{aligned}$$

Now setting $v = x = 0$, we obtain

$$au - by = ug(tu, sy) - yh(tu, sy), \quad (4.91)$$

and finally replacing u by $\frac{u}{t}$ and y by $\frac{y}{s}$ and multiplying by ts yields

$$asu - bty = sug(u, y) - tyh(u, y). \quad (4.92)$$

Replacing u by x in (4.92) and adding the result to t times (4.89) gives

$$a(s+t)x = (s+t)xg(x, y), \quad (4.93)$$

and since $s^2 \neq t^2$ we have

$$g(x, y) = a, \quad (4.94)$$

for all $x \in \mathbb{R} \setminus \{0\}$ and $y \in \mathbb{R}$. A similar argument yields $h(x, y) = b$ for all $x \in \mathbb{R}$ and $y \in \mathbb{R} \setminus \{0\}$. Letting $u = 1 + x$ and $v = 1 + y$ in (4.46) and then using (4.90), we have

$$a + b = g(t + x(s+t), t + y(s+t)) + h(t + x(s+t), t + y(s+t))$$

which indeed yields

$$a + b = g(x, y) + h(x, y)$$

for all $x, y \in \mathbb{R}$. Now substituting $x = 0$ in above, we have $a + b = g(0, y) + h(0, y)$. Since $h(x, y) = b$ for all $x \in \mathbb{R}$ and $y \in \mathbb{R} \setminus \{0\}$, we get $g(0, y) = a$ for all $y \in \mathbb{R} \setminus \{0\}$. Further, since $g(0, 0) = a$, we see that (4.94) holds for all $x, y \in \mathbb{R}$.

Substituting (4.94) and (4.90) into (4.46), yields

$$h(x, y) = b, \quad (4.95)$$

for all $x, y \in \mathbb{R}$; putting it all together we get

$$\left. \begin{array}{rcl} f(x, y) & = & ax + by + c \\ g(x, y) & = & a \\ h(x, y) & = & b \end{array} \right\} \quad (4.96)$$

for all $x, y \in \mathbb{R}$, where a, b, c are arbitrary real constants.

The following is an easy consequence of the Theorem 4.4.

Corollary 4.1 *Let f_x and f_y be the partial derivatives of $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, and let s and t , a priori chosen real parameters. The function f satisfies the functional-differential equation*

$$\begin{aligned} & f(u, v) - f(x, y) \\ &= (u - x)f_x(su + tx, sv + ty) + (v - y)f_y(su + tx, sv + ty), \end{aligned}$$

for all $x, y, u, v \in \mathbb{R}$ with $(u - x)^2 + (v - y)^2 \neq 0$, if and only if

$$f(x, y) = \begin{cases} ax^2 + bx + cy^2 + dy + exy + \alpha & \text{if } s = \frac{1}{2} = t \\ ax + by + c & \text{otherwise,} \end{cases}$$

where a, b, c, d, e, α are arbitrary real constants.

4.4 Cauchy's MVT for Functions in Two Variables

The mean value theorem for functions in two variables can be generalized to include two functions as follows:

Theorem 4.5 *For all real valued functions g and f in \mathbb{R}^2 with continuous partial derivatives f_x , f_y , g_x and g_y and for all distinct pairs (x, y) and (u, v) in \mathbb{R}^2 , there exists an intermediate point (η, ξ) on the line segment joining the points (x, y) and (u, v) such that*

$$\begin{aligned} & [f(u, v) - f(x, y)] [(u - x)g_x(\eta, \xi) + (v - y)g_y(\eta, \xi)] \\ = & [g(u, v) - g(x, y)] [(u - x)f_x(\eta, \xi) + (v - y)f_y(\eta, \xi)]. \end{aligned}$$

Proof: The proof is analogous to the one for the one variable case. We define an auxiliary function

$$\Psi(s, t) = [f(u, v) - f(s, t)][g(u, v) - g(x, y)] - [f(u, v) - f(x, y)][g(u, v) - g(s, t)],$$

then $\Psi(u, v) = \Psi(x, y) = 0$, Ψ is differentiable wherever f and g are, hence by the mean value theorem for functions of two variables there is a point (η, ξ) on the line segment joining (x, y) and (u, v) , such that

$$(u - x)\Psi_s(\eta, \xi) + (v - y)\Psi_t(\eta, \xi) = 0.$$

Carrying out the partial derivatives on Ψ yields

$$\begin{aligned} & (u - x)\{g_x(\eta, \xi)[f(u, v) - f(x, y)] - f_x(\eta, \xi)[g(u, v) - g(x, y)]\} \\ + & (v - y)\{f_y(\eta, \xi)[g(u, v) - g(x, y)] - g_y(\eta, \xi)[f(u, v) - f(x, y)]\} = 0 \end{aligned}$$

which proves the result.

4.5 Some Open Problems

We conclude this chapter with the following problems. The equation that appears in the above theorem yields a functional equation for given η and ξ . Assuming $\eta(x, u) = x + u$ and $\xi(v, y) = v + y$, and replacing the partial derivatives of f and g by unknown functions h, k, ℓ, m in the above equation, one obtains

$$\begin{aligned} & [f(u, v) - f(x, y)] [(u - x)h(u + x, v + y) + (v - y)k(u + x, v + y)] \\ = & [g(u, v) - g(x, y)] [(u - x)\ell(u + x, v + y) + (v - y)m(u + x, v + y)] \end{aligned}$$

for all $x, y, u, v \in \mathbb{R}$ with $(u - x)^2 + (v - y)^2 \neq 0$. The general solution of this functional equation is not known. Further, the functional equation (4.5) can be generalized to

$$\begin{aligned} & f(u, v) - f(x, y) \\ &= [\ell(u) - \ell(x)] g(u + x, v + y) + [\ell(v) - \ell(y)] h(u + x, v + y), \end{aligned} \quad (4.97)$$

for all $x, y, u, v \in \mathbb{R}$ with $(u - x)^2 + (v - y)^2 \neq 0$. Here $\ell : \mathbb{R} \rightarrow \mathbb{R}$ is a continuous strictly monotone function. This equation also needs an investigation.

We have not been able to generalize Pompeiu's mean value theorem for functions in two variables and determine a corresponding functional equation.

The equation (4.5) characterizes quadratic polynomials of at most degree two in two variables. It is not clear what the generalization of (4.5) should be for higher order polynomials in x and y . The idea of divided difference does not lead us to any meaningful generalization of the functional equation (4.5) for characterizing the higher order polynomials.

Chapter 5

Some Generalizations of Lagrange's Mean Value Theorem

In this chapter, we discuss several generalizations of Lagrange's mean value theorem. In section one, we examine mean value theorems due to Flett (1958) and Trahan (1966). This section is devoted to generalizations of the mean value theorem for real valued functions on the reals. In section two, we study some generalizations of Lagrange's mean value theorem and Flett's mean value theorem for real valued functions on \mathbb{R}^2 . In this section, we also present some results by Clarke and Ledyaev (1994). Section three is devoted to mean value theorems for vector valued functions on the reals. Some results of Sanderson (1972) and McLeod (1964) are discussed in this section. Section four addresses the various generalizations of the mean value theorem for vector valued functions on the plane. We also present some results due to Furi and Martelli (1995) in this section and examine some consequences to illustrate the generality and simplicity of their result. In section five, we discuss the generalizations of the mean value theorem for complex-valued functions on the complex plane. Among others, we study a generalization due to Evard and Jafari (1992). A local version of Lagrange's mean value theorem for holomorphic functions, due to Samuelsson (1973), is also treated here. In section six, we discuss a conjecture due to Furi and Martelli (1995) regarding a multidimensional version of Rolle's theorem.

5.1 MVTs for Real Functions

In chapter two, we examined Lagrange's mean value theorem which says that for every real-valued function $f : [a, b] \rightarrow \mathbb{R}$, continuous on $[a, b]$ and

differentiable on the open interval $]a, b[$, there exists a $\eta \in]a, b[$ such that

$$f(b) - f(a) = (b - a) f'(\eta). \quad (5.1)$$

This mean value theorem was derived using Rolle's theorem which states that for every real-valued function $g : [a, b] \rightarrow \mathbb{R}$ with $g(a) = g(b)$, continuous on $[a, b]$ and differentiable on the open interval $]a, b[$ there exists a $\eta \in]a, b[$ such that

$$g'(\eta) = 0. \quad (5.2)$$

If we replace the function f in the mean value theorem by another function g defined as

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a} x, \quad (5.3)$$

then $g(a) = g(b)$ and the two results are now equivalent. Hence, from here after we shall not distinguish between Rolle's theorem and Lagrange's mean value theorem, and consider them both as the mean value theorem. Let us examine how the mean value theorem is deduced from Rolle's theorem. The usual proof of the mean value theorem is to apply the Rolle's theorem to functions suitably designed to yield the desired result. Frequently, no mention is made of how these functions are discovered.

Notice that (5.1) can be represented as a determinant, namely

$$f(b) - f(a) - (b - a) f'(\eta) = \begin{vmatrix} 1 & f'(\eta) & 0 \\ b & f(b) & 1 \\ a & f(a) & 1 \end{vmatrix}.$$

Similarly, the auxiliary function,

$$g(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a} (x - a)$$

used for deducing the mean value theorem can also be represented as a determinant, namely

$$g(x) = -\frac{1}{b - a} \begin{vmatrix} x & f(x) & 1 \\ b & f(b) & 1 \\ a & f(a) & 1 \end{vmatrix}.$$

It is easy to see that one may use the above determinant instead of g and can deduce the same result. From analytic geometry we know that the above

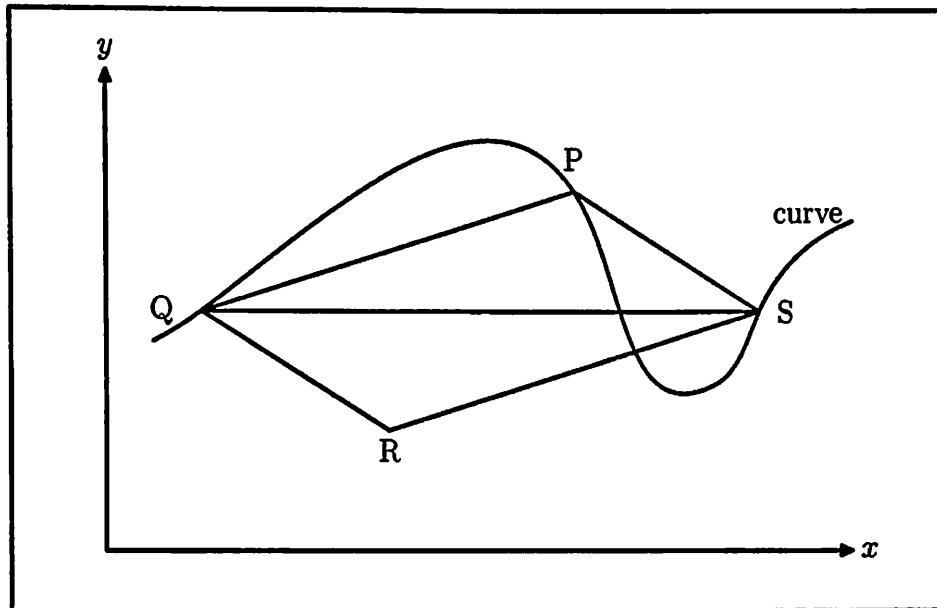


Fig. 5.1 A Geometrical Illustration of the Auxiliary Function.

determinant represents the area of the parallelogram $PQRS$ (see Figure 5.1). This interpretation is very insightful and allows one to generalize the mean value theorem even further. For one such generalization of the mean value theorem see Barrett and Jacobson (1960). We present two consequences of the mean value theorem not included in most textbooks. These two results appeared in Furi and Martelli (1991).

Lemma 5.1 *Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous on $[a, b]$ and differentiable on the open interval $]a, b[$ except possibly at finitely many points. Then there exists a point $c \in]a, b[$ such that*

$$|f(b) - f(a)| \leq (b - a) |f'(c)|. \quad (5.4)$$

Proof: Assume that there is only one point $d \in]a, b[$ where the function f is not differentiable. By applying the mean value theorem to f on $[a, d]$ and $[d, b]$ respectively, we obtain

$$f(d) - f(a) = (d - a) f'(c_1),$$

and

$$f(b) - f(d) = (b - d) f'(c_2),$$

for some $c_1 \in]a, d[$ and $c_2 \in]d, b[$. Adding the above two, we get

$$f(b) - f(a) = (d - a) f'(c_1) + (b - d) f'(c_2)$$

and from which we get

$$\begin{aligned} |f(b) - f(a)| &\leq (d - a) |f'(c_1)| + (b - d) |f'(c_2)| \\ &\leq (d - a) |f'(c)| + (b - d) |f'(c)| \\ &= (b - a) |f'(c)|, \end{aligned}$$

where

$$|f'(c)| = \max \{|f'(c_1)|, |f'(c_2)|\}.$$

The proof can obviously be extended to the case when f is not differentiable at more than one point.

Lemma 5.2 *Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous on $[a, b]$ and differentiable on the open interval $]a, b[$ except possibly at a finite number, n , of points. Then there exist $n+1$ points $c_1, c_2, \dots, c_{n+1} \in]a, b[$ and $n+1$ positive numbers $\alpha_1, \alpha_2, \dots, \alpha_{n+1}$ such that*

$$\alpha_1 + \alpha_2 + \cdots + \alpha_{n+1} = 1$$

and

$$f(b) - f(a) = (b - a) \sum_{i=1}^{n+1} \alpha_i f'(c_i). \quad (5.5)$$

Proof: Assume that there is only one point $d \in]a, b[$ where the function f is not differentiable. By applying the mean value theorem to f on $[a, d]$ and $[d, b]$ respectively, we get

$$f(d) - f(a) = (d - a) f'(c_1),$$

and

$$f(b) - f(d) = (b - d) f'(c_2),$$

for some $c_1 \in]a, d[$ and $c_2 \in]d, b[$. Adding the two equations, we get

$$f(b) - f(a) = (d - a) f'(c_1) + (b - d) f'(c_2).$$

Rewriting this, we obtain

$$f(b) - f(a) = \left[\frac{d-a}{b-a} f'(c_1) + \frac{b-d}{b-a} f'(c_2) \right] (b-a)$$

which is

$$f(b) - f(a) = [\alpha_1 f'(c_1) + \alpha_2 f'(c_2)] (b-a),$$

where

$$\alpha_1 = \frac{d-a}{b-a} \quad \text{and} \quad \alpha_2 = \frac{b-d}{b-a}.$$

Clearly $\alpha_1 + \alpha_2 = 1$, $\alpha_1 > 0$ and $\alpha_2 > 0$. This proof can obviously be extended to the case when f is not differentiable at $n > 1$ points.

In 1958, T.M. Flett proved the following result which is a variant of Lagrange's mean value theorem.

Theorem 5.1 *Let $f : [a, b] \rightarrow \mathbb{R}$ be differentiable on $[a, b]$ and $f'(a) = f'(b)$. Then there exists a point $\eta \in]a, b[$ such that*

$$f(\eta) - f(a) = (\eta - a) f'(\eta). \quad (5.6)$$

Proof: Without loss of generality, we shall assume that $f'(a) = f'(b) = 0$. If this is not the case we work with $f(x) - x f'(a)$. Consider the function $g : [a, b] \rightarrow \mathbb{R}$ defined by

$$g(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{if } x \in]a, b] \\ f'(a) & \text{if } x = a. \end{cases} \quad (5.7)$$

Evidently g is continuous on $[a, b]$ and differentiable on $]a, b[$. Further, from (5.7) we have

$$g'(x) = -\frac{f(x) - f(a)}{(x-a)^2} + \frac{f'(x)}{x-a}$$

which is

$$g'(x) = -\frac{g(x)}{x-a} + \frac{f'(x)}{x-a} \quad (5.8)$$

for all $x \in]a, b[$. In view of (5.7), to establish the theorem we have to show that there exists a point $\eta \in]a, b[$ such that $g'(\eta) = 0$.

From (5.7), we see that $g(a) = 0$. If $g(b) = 0$, then by Rolle's theorem there exists an $\eta \in]a, b[$ such that $g'(\eta) = 0$ and theorem is established. If

$g(b) \neq 0$, then either $g(b) > 0$ or $g(b) < 0$. Suppose $g(b) > 0$. Then from (5.8), we see that

$$g'(b) = -\frac{g(b)}{b-a} < 0.$$

Since g is continuous and $g'(b) < 0$, there exists a point x_1 in $]a, b[$ such that

$$g(x_1) > g(b).$$

Hence, we have $g(a) < g(b) < g(x_1)$ and by intermediate value theorem there exists a $x_o \in]a, x_1[$ such that $g(x_o) = g(b)$. Now applying the Rolle's theorem to the function g on the interval $[x_o, b]$ we have $g'(\eta) = 0$ for some $\eta \in]a, b[$.

A similar argument applies if $g(b) < 0$, and now the proof of the theorem is complete.

The geometrical interpretation of this theorem is the following. If the curve $y = f(x)$ has a continuously turning tangent in $a < x < b$, and if the tangents at $x = a$ and $x = b$ are parallel, then there is an intermediate point η such that the tangent there passes through the point a . The Figure 5.2 geometrically illustrates the mean value theorem of Flett.

The following theorem removes the boundary assumption on the derivative of f , that is $f'(a) = f'(b)$.

Theorem 5.2 *If $f : [a, b] \rightarrow \mathbb{R}$ is a differentiable function, then there exists a point $\eta \in]a, b[$ such that*

$$f(\eta) - f(a) = (\eta - a) f'(\eta) - \frac{1}{2} \frac{f'(b) - f'(a)}{b - a} (\eta - a)^2.$$

Proof: Defining an auxiliary function $\psi : [a, b] \rightarrow \mathbb{R}$ as

$$\psi(x) = f(x) - \frac{1}{2} \frac{f'(b) - f'(a)}{b - a} (x - a)^2$$

we see that ψ is differentiable on $[a, b]$ and

$$\psi'(x) = f'(x) - \frac{f'(b) - f'(a)}{b - a} (x - a).$$

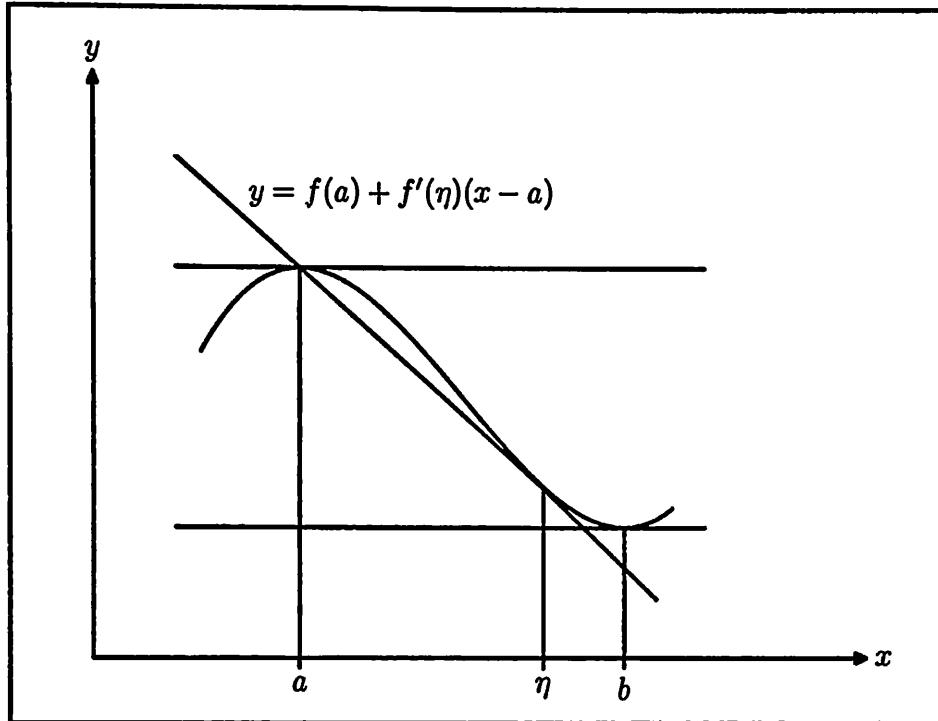


Fig. 5.2 A Geometrical Illustration of Flett's Theorem.

From this it is easy to check that $\psi'(a) = \psi'(b) = f'(a)$. Applying Flett's mean value theorem to ψ , we get

$$\psi(\eta) - \psi(a) = (\eta - a) \psi'(\eta)$$

for some $\eta \in]a, b[$. Using the definition of the auxiliary function, we get the asserted result

$$f(\eta) - f(a) = (\eta - a) f'(\eta) - \frac{1}{2} \frac{f'(b) - f'(a)}{b - a} (\eta - a)^2.$$

This completes the proof of the theorem.

The auxiliary function ψ used in the above theorem is obtained by considering the difference between $f(x)$ and a quadratic approximation, $A + B(x-a) + C(x-a)^2$, of $f(x)$ and then imposing the boundary condition on the derivative of ψ , namely $\psi'(a) = \psi'(b)$. The boundary condition on the derivative of ψ yields $C = \frac{1}{2} \frac{f'(b) - f'(a)}{b - a}$. The constants A and B are arbitrary and we have chosen them to be zero for the sake of convenience.

In the next theorem, we present the generalization of Flett's theorem due to D.H. Trahan (1966). To prove his result, we need two basic lemmas.

Lemma 5.3 *Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous on $[a, b]$, differentiable on $]a, b]$ and $[f(b) - f(a)] f'(b) \leq 0$. Then there exists a point $\eta \in]a, b]$ such that $f'(\eta) = 0$.*

Proof: If $f(b) = f(a)$, then by Rolle's theorem there exists a point $\eta \in]a, b[$ such that $f'(\eta) = 0$. If $f'(b) = 0$, then letting $\eta = b$ we have $f'(\eta) = 0$. Next, suppose $[f(b) - f(a)] f'(b) < 0$. This implies that either $f'(b) < 0$ and $f(b) > f(a)$ or $f'(b) > 0$ and $f(b) < f(a)$. In the first case, since f is continuous on $[a, b]$ and $f(b) > f(a)$ with f decreasing at b , the function f has a maximum at $\eta \in]a, b[$. Hence $f'(\eta) = 0$. Similarly, in the second case f has a minimum at some point $\eta \in]a, b[$ and hence $f'(\eta) = 0$. The proof of the lemma is now complete.

The following lemma is obvious from the one above.

Lemma 5.4 *Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous on $[a, b]$, differentiable on $]a, b]$ and $[f(b) - f(a)] f'(b) < 0$. Then there exists a point $\eta \in]a, b[$ such that $f'(\eta) = 0$.*

Now we present a generalization of Flett's mean value theorem.

Theorem 5.3 *Let $f : [a, b] \rightarrow \mathbb{R}$ be differentiable on $[a, b]$ and*

$$\left[f'(b) - \frac{f(b) - f(a)}{b - a} \right] \left[f'(a) - \frac{f(b) - f(a)}{b - a} \right] \geq 0. \quad (5.9)$$

Then there exists a point $\eta \in]a, b]$ such that

$$f(\eta) - f(a) = (\eta - a) f'(\eta).$$

Proof: Let us define a function $h : [a, b] \rightarrow \mathbb{R}$ by

$$h(x) = \begin{cases} \frac{f(x) - f(a)}{x - a} & \text{if } x \in]a, b] \\ f'(a) & \text{if } x = a. \end{cases} \quad (5.10)$$

Then h is continuous on the interval $[a, b]$ and it is differentiable on $]a, b]$. Differentiating h , we get

$$h'(x) = -\frac{f(x) - f(a)}{(x - a)^2} + \frac{f'(x)}{x - a}$$

for all $x \in]a, b]$. Since

$$[h(b) - h(a)] h'(b) = \frac{-1}{b-a} \left[f'(b) - \frac{f(b) - f(a)}{b-a} \right] \left[f'(a) - \frac{f(b) - f(a)}{b-a} \right],$$

by (5.9), we see that

$$[h(b) - h(a)] h'(b) \leq 0.$$

Hence, applying Lemma 5.4 we obtain

$$h'(\eta) = 0$$

for some $\eta \in]a, b]$. By the definition of h this amounts to

$$f(\eta) - f(a) = (\eta - a) f'(\eta)$$

and the proof of the theorem is now complete.

Our Theorem 5.3 is a generalization of Theorem 5.1 since the latter can be deduced from the former. To see this, define h as in the Theorem 5.3. That is

$$h(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{if } x \in]a, b] \\ f'(a) & \text{if } x = a. \end{cases} \quad (5.11)$$

Then h is continuous on the interval $[a, b]$ and differentiable on $]a, b]$. Differentiating h , we get

$$h'(x) = -\frac{f(x) - f(a)}{(x-a)^2} + \frac{f'(x)}{x-a}$$

for all x in $a < x \leq b$. First, consider the case when

$$f(b) - f(a) = (b-a) f'(b). \quad (5.12)$$

Using (5.11) and (5.12), we have

$$\begin{aligned} h(b) - h(a) &= \frac{f(b) - f(a)}{b-a} - f'(a) && \text{by (5.11)} \\ &= f'(b) - f'(a) && \text{by (5.12)} \\ &= 0 && \text{by hypothesis} \end{aligned}$$

Hence, we have $h(b) = h(a)$. Applying Rolle's theorem to h , we get $h'(\eta) = 0$ for some $\eta \in]a, b[$. This yields

$$f(\eta) - f(a) = (\eta - a) f'(\eta).$$

Next, we consider the case

$$f(b) - f(a) \neq (b - a) f'(b). \quad (5.13)$$

Hence, we have $\left[f'(b) - \frac{f(b) - f(a)}{b - a} \right] > 0$ or $\left[f'(b) - \frac{f(b) - f(a)}{b - a} \right] < 0$. Therefore, using the fact that $f'(b) = f'(a)$, we get

$$\left[f'(b) - \frac{f(b) - f(a)}{b - a} \right] \left[f'(a) - \frac{f(b) - f(a)}{b - a} \right] > 0.$$

Thus, using Theorem 5.3 we have Theorem 5.1. It can be shown that $\eta \neq b$.

5.2 MVTs for Real Valued Functions on the Plane

In Chapter 4, we treated the following generalization of Lagrange's mean value theorem for real valued functions in two variables. For every function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ with continuous partial derivatives f_x and f_y and for all distinct pairs (x_1, x_2) and (y_1, y_2) in \mathbb{R}^2 , there exists an intermediate point (η_1, η_2) on the line segment joining the points (x_1, x_2) and (y_1, y_2) such that

$$f(y_1, y_2) - f(x_1, x_2) = (y_1 - x_1) f_x(\eta_1, \eta_2) + (y_2 - x_2) f_y(\eta_1, \eta_2). \quad (5.14)$$

Definition 5.1 Let (x_1, x_2) and (y_1, y_2) be any two points in \mathbb{R}^2 . The *Euclidean inner product* $\langle (x_1, x_2), (y_1, y_2) \rangle$ between the points (x_1, x_2) and (y_1, y_2) is defined as

$$\langle (x_1, x_2), (y_1, y_2) \rangle = \sum_{i=1}^2 x_i y_i,$$

and the *norm* $\|(x_1, x_2)\|$ of (x_1, x_2) is defined as

$$\|(x_1, x_2)\| = \sqrt{\langle (x_1, x_2), (x_1, x_2) \rangle}.$$

For notational simplicity, if $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a function in two variables and f_x and f_y are its partial derivatives, then we denote (f_x, f_y) as f' .

Using the above definition and notation, we can rewrite (5.14) as

$$f(y_1, y_2) - f(x_1, x_2) = \langle f'(\eta_1, \eta_2), (y_1 - x_1, y_2 - x_2) \rangle.$$

This can be further simplified to

$$f(y) - f(x) = \langle f'(\eta), y - x \rangle$$

where $x = (x_1, x_2)$, $y = (y_1, y_2)$ and $\eta = (\eta_1, \eta_2)$.

Now we present a generalization of Flett's mean value theorem for real valued functions in two variables.

Theorem 5.4 *For every function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ with continuous partial derivatives f_x and f_y and for all distinct pairs a and b in \mathbb{R}^2 with $f'(a) = f'(b)$, there exists an intermediate point η on the line segment joining a and b such that*

$$f(\eta) - f(a) = \langle \eta - a, f'(\eta) \rangle. \quad (5.15)$$

Proof: Let $a = (a_1, a_2)$ and $b = (b_1, b_2)$ be any two points on the plane \mathbb{R}^2 . Let $h_1 = b_1 - a_1$ and $h_2 = b_2 - a_2$. Let \mathcal{L} be the line segment obtained by joining the points a and b . The co-ordinates of any point on this line segment is given by $(a_1 + h_1t, a_2 + h_2t)$ for some $t \in [0, 1]$. We define a function $F : [0, 1] \rightarrow \mathbb{R}$ by

$$F(t) = f(a_1 + h_1t, a_2 + h_2t)$$

keeping a and b fixed temporarily. Evidently, F is differentiable on $[0, 1]$. The derivative of this function is

$$F'(t) = h_1 f_x(a_1 + h_1t, a_2 + h_2t) + h_2 f_y(a_1 + h_1t, a_2 + h_2t).$$

Since $f'(a) = f'(b)$, we see that $F'(0) = F'(1)$. Applying Flett's mean value theorem to F , we get

$$\frac{F(t_o) - F(0)}{t_o - 0} = F'(t_o),$$

for some $t_o \in]0, 1[$. Now applying the definition of F to this, we get

$$\begin{aligned} & f(a_1 + h_1 t_o, a_2 + h_2 t_o) - f(a_1, a_2) \\ &= t_o h_1 f_x(a_1 + h_1 t_o, a_2 + h_2 t_o) + t_o h_2 f_y(a_1 + h_1 t_o, a_2 + h_2 t_o). \end{aligned}$$

From this, we have

$$f(\eta_1, \eta_2) - f(a_1, a_2) = (\eta_1 - a_1) f_x(\eta_1, \eta_2) + (\eta_2 - a_2) f_y(\eta_1, \eta_2)$$

which is

$$f(\eta) - f(a) = \langle \eta - a, f'(\eta) \rangle.$$

This completes the proof of the theorem.

We improve the above theorem similar to an improvement provided by Trahan (1966) for functions in one variable.

Theorem 5.5 *For every function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ with continuous partial derivatives f_x and f_y and for all distinct pairs a and b in \mathbb{R}^2 with*

$$[\langle b - a, f'(a) \rangle - f(b) + f(a)] [\langle b - a, f'(b) \rangle - f(b) + f(a)] \geq 0, \quad (5.16)$$

there exists an intermediate point η on the line segment joining a and b such that

$$f(\eta) - f(a) = \langle \eta - a, f'(\eta) \rangle.$$

Proof: Let $a = (a_1, a_2)$ and $b = (b_1, b_2)$ be any two points on the plane \mathbb{R}^2 . Let $h_1 = b_1 - a_1$ and $h_2 = b_2 - a_2$. Let \mathcal{L} be the line segment obtained by joining the points a and b . The co-ordinates of any point on this line segment \mathcal{L} is given by

$$(a_1 + h_1 t, a_2 + h_2 t)$$

for some $t \in [0, 1]$. We define a function $F : [0, 1] \rightarrow \mathbb{R}$ by

$$F(t) = f(a_1 + h_1 t, a_2 + h_2 t) \quad (5.17)$$

keeping a and b fixed temporarily. Evidently, F is differentiable on $[0, 1]$. The derivative of this function is

$$F'(t) = h_1 f_x(a_1 + h_1 t, a_2 + h_2 t) + h_2 f_y(a_1 + h_1 t, a_2 + h_2 t). \quad (5.18)$$

From (5.18), we obtain

$$F'(0) = (b_1 - a_1) f_x(a_1, a_2) + (b_2 - a_2) f_y(a_1, a_2) = \langle b - a, f'(a) \rangle. \quad (5.19)$$

Similarly, we have

$$F'(1) = (b_1 - a_1) f_x(b_1, b_2) + (b_2 - a_2) f_y(b_1, b_2) = \langle b - a, f'(b) \rangle. \quad (5.20)$$

Using (5.17), (5.19) and (5.20) in (5.16), we see that

$$[F'(0) - F(1) + F(0)] [F'(1) - F(1) + F(0)] \geq 0.$$

Applying Theorem 5.3 to F on the interval $[0, 1]$, we get

$$\frac{F(t_o) - F(0)}{t_o - 0} = F'(t_o), \quad (5.21)$$

for some $t_o \in]0, 1[$. Writing $\eta = (\eta_1, \eta_2)$, where $\eta_i = a_i + h_i t_o$ ($i = 1, 2$) we see that

$$\begin{aligned} t_o F'(t_o) &= t_o h_1 f_x(\eta_1, \eta_2) + t_o h_2 f_y(\eta_1, \eta_2) \\ &= (\eta_1 - a_1) f_x(\eta_1, \eta_2) + (\eta_2 - a_2) f_y(\eta_1, \eta_2) \\ &= \langle \eta - a, f'(\eta) \rangle. \end{aligned}$$

From (5.21) by (5.17) and the above relation, we obtain

$$f(\eta) - f(a) = \langle \eta - a, f'(\eta) \rangle$$

and the proof of the theorem is now complete.

Definition 5.2 Let X and Y be any two subsets of \mathbb{R}^2 . Then the set

$$\{x + (y - x)t \mid x \in X, y \in Y, t \in [0, 1]\}$$

is said to be a *segment* in \mathbb{R}^2 and is denoted by $[X, Y]$.

Note that if X and Y are convex and compact sets, then $[X, Y]$ is also a convex and compact set in \mathbb{R}^2 . In fact, $[X, Y]$ is the convex hull of the set $X \cup Y$.

Now we present a result due to Clarke and Ledyaev (1993).

Theorem 5.6 Let X and Y be convex and compact sets in \mathbb{R}^2 , and let a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be continuously differentiable on $[X, Y]$. Then there exists a point $\eta \in [X, Y]$ such that

$$\min_{y \in Y} f(y) - \max_{x \in X} f(x) \leq \langle y - x, f'(\eta) \rangle \quad (5.22)$$

holds for all $x \in X$ and $y \in Y$.

Proof: Let $x \in X$ and $y \in Y$ be any two points. We apply the mean value theorem for real valued functions in two variables to the function f to get

$$f(y) - f(x) = \langle y - x, f'(\eta_o) \rangle$$

for some η_o in the line segment joining x and y . This equality implies the following inequality

$$\min_{y \in Y} f(y) - \max_{x \in X} f(x) \leq \langle y - x, f'(\eta_o) \rangle.$$

The function $\phi : [X, Y] \rightarrow \mathbb{R}$ defined by

$$\phi(\xi) = \langle y - x, f'(\xi) \rangle$$

for temporarily fixed x and y is continuous in $[X, Y]$ since f is a continuously differentiable function on $[X, Y]$. Hence f has a maximum on the set $[X, Y]$. Let this maximum be achieved at the point η . Hence, we have

$$\phi(\xi) \leq \phi(\eta) \quad \text{for all } \xi \in [X, Y].$$

From this, we see that

$$\min_{y \in Y} f(y) - \max_{x \in X} f(x) \leq \langle y - x, f'(\eta) \rangle$$

holds for all $x \in X$ and $y \in Y$. This completes the proof.

The convexity of the compact sets X and Y is essential. Clarke and Ledyayev (1993) wrote that it is yet unclear whether in the statement of the theorem it is possible to replace the continuous differentiability of f by its differentiability.

5.3 MVTs for Vector Valued Functions on the Reals

Now we turn to the problem of finding a generalization of Lagrange's mean value theorem for differentiable functions on \mathbb{R} to \mathbb{R}^2 . The straightforward generalization of the mean value theorem for functions from \mathbb{R} into \mathbb{R}^2 fails. The mean value theorem states that if a differentiable real function f is constant on the boundary of an interval $[a, b]$ (that is, $f(a) = f(b)$), then it has a critical point in the interval $]a, b[$. Consider the function $f : \mathbb{R} \rightarrow \mathbb{R}^2$ defined by

$$f(x) = (x - x^2, x - x^3).$$

The derivative of the function f is given by

$$f'(x) = (1 - 2x, 1 - 3x^2).$$

This function is constant on the boundary of the interval $[0, 1]$, that is

$$f(0) = (0, 0) \quad \text{and} \quad f(1) = (0, 0).$$

However, the function $f(x)$ has no critical point in $]0, 1[$, that is the derivative $f'(x)$ is never zero in $]0, 1[$. Therefore, the mean value theorem does not extend in a straightforward manner to 2-dimensional vector valued functions on the reals.

We have seen that Rolle's theorem is equivalent to the mean value theorem for functions from the reals into the reals. In this section and in the remaining sections of this chapter we will focus on Rolle's theorem rather than the mean value theorem. The reason for this is that Rolle's theorem has a natural extension in higher dimensions. In this section, we present a generalization due to Sanderson (1972).

Theorem 5.7 *Let $f : [a, b] \rightarrow \mathbb{R}^2$ be a differentiable 2-dimensional vector valued function and $f(a), f(b)$ be orthogonal to a non-zero vector v in \mathbb{R}^2 . Then there exists a point $\eta \in]a, b[$ such that*

$$\langle v, f'(\eta) \rangle = 0,$$

that is, $f'(\eta)$ is orthogonal to the vector $v \in \mathbb{R}^2$.

Proof: First, we define a function $F : [a, b] \rightarrow \mathbb{R}$ by

$$F(t) = \langle v, f(t) \rangle.$$

Thus F is differentiable on $[a, b]$ and $F(a) = F(b) = 0$. Applying Rolle's Theorem to F , we get

$$F'(\eta) = 0$$

for some $\eta \in]a, b[$. This amounts to

$$\langle v, f'(\eta) \rangle = 0$$

by definition of F and the proof is now complete.

From this theorem, one can deduce the mean value theorem in the following manner. If we take $f(t) = (t, g(t))$ for t in the interval $[a, b]$ and $v = (g(b) - g(a), a - b)$, then applying the above theorem, we get

$$\langle v, f'(\eta) \rangle = 0$$

for some $\eta \in]a, b[$. From this, we obtain

$$\langle (1, g'(\eta)), (g(b) - g(a), a - b) \rangle = 0$$

which yields

$$g(b) - g(a) = (b - a) g'(\eta).$$

Similarly, Cauchy's mean value theorem can be deduced from the above theorem by letting $f(t) = (h(t), g(t))$ and $v = (g(b) - g(a), h(a) - h(b))$.

The following theorem is a variation of the above theorem.

Theorem 5.8 *Let $f : [a, b] \rightarrow \mathbb{R}^2$ be a differentiable 2-dimensional vector valued function which is orthogonal to a non-zero vector v in \mathbb{R}^2 at two distinct points of $[a, b]$. Then there exists a point $\eta \in]a, b[$ such that*

$$\langle v, f'(\eta) \rangle = 0,$$

that is, $f'(\eta)$ is orthogonal to the vector $v \in \mathbb{R}^2$.

Proof: Let c and d be two distinct points in $[a, b]$. We apply the above theorem to f on the interval $[c, d]$. Hence, we conclude that there is a point $\eta \in]c, d[$ such that the nonzero vector v is orthogonal to $f'(\eta)$, and the proof of the theorem is now complete.

The above two theorems can be further generalized to the followings theorems. We will leave their proof to the reader.

Theorem 5.9 *Let $f : [a, b] \rightarrow \mathbb{R}^n$ be a differentiable n -dimensional vector valued function and $f(a), f(b)$ and the first $k - 1$ derivatives of f at a be orthogonal to a non-zero vector v in \mathbb{R}^n . Then there exists a point $\eta \in]a, b[$ such that*

$$\langle v, f^{(k)}(\eta) \rangle = 0,$$

that is, $f^{(k)}(\eta)$ is orthogonal to the vector $v \in \mathbb{R}^n$.

Theorem 5.10 *Let $f : [a, b] \rightarrow \mathbb{R}^n$ be a differentiable n -dimensional vector valued function which is orthogonal to a non-zero vector v in \mathbb{R}^n at $k + 1$ distinct points of $[a, b]$. Then there exists a point $\eta \in]a, b[$ such that*

$$\langle v, f^{(k)}(\eta) \rangle = 0,$$

that is, $f^{(k)}(\eta)$ is orthogonal to the vector $v \in \mathbb{R}^n$.

The following theorem is due to McLeod (1964) whose proof is beyond the scope of this introductory book.

Theorem 5.11 *Let $f : [a, b] \rightarrow \mathbb{R}^n$ be continuous on $[a, b]$ differentiable on $]a, b[$. Further assume that $f'(x)$ is continuous from the right. Then there are n points $c_1, c_2, \dots, c_n \in]a, b[$ and n positive numbers r_1, r_2, \dots, r_n such that*

$$r_1 + r_2 + \dots + r_n = 1$$

and

$$f(b) - f(a) = (b - a) \sum_{k=1}^n r_k f'(c_k).$$

Note that if $n = 1$, then the above theorem reduces to the well known Lagrange's mean value theorem.

5.4 MVTs for Vector Valued Functions on the Plane

We have seen that for vector valued functions on the reals there is no straightforward generalization of the mean value theorem. In this section, we will illustrate with an example that there is no straightforward generalization of the mean value theorem for vector valued functions on the plane either.

A function f from \mathbb{R}^2 into \mathbb{R}^2 is a pair of functions

$$f : \begin{cases} u = f_1(x, y) \\ v = f_2(x, y), \end{cases}$$

where $f_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$. The functions f_1 and f_2 are called the co-ordinate functions of f . Hence f can be written as

$$(u, v) = f(x, y) = (f_1(x, y), f_2(x, y)).$$

A function $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is said to be differentiable at a point (x, y) of \mathbb{R}^2 if each of the co-ordinate functions f_1 and f_2 is differentiable at (x, y) . The derivative of f at (x, y) is given by the matrix

$$\begin{pmatrix} \frac{\partial f_1(x,y)}{\partial x} & \frac{\partial f_1(x,y)}{\partial y} \\ \frac{\partial f_2(x,y)}{\partial x} & \frac{\partial f_2(x,y)}{\partial y} \end{pmatrix}$$

and denoted by $f'(x, y)$. This matrix is also known as the Jacobian of the function $f(x, y)$. By

$$\begin{pmatrix} \frac{\partial f_1(\eta,\xi)}{\partial x} & \frac{\partial f_1(\eta,\xi)}{\partial y} \\ \frac{\partial f_2(\eta,\xi)}{\partial x} & \frac{\partial f_2(\eta,\xi)}{\partial y} \end{pmatrix}$$

we mean the Jacobian of $f(x, y)$ evaluated at the point $(\eta, \xi) \in \mathbb{R}^2$.

We need the following terminologies to state precisely the multidimensional version of Rolle's theorem. Let $D_r(x_o, y_o)$ denote the ball of radius r ,

that is the set $\{(x, y) \in \mathbb{R}^2 \mid \|(x - x_o, y - y_o)\| \leq r\}$. Similarly, let $B_r(x_o, y_o)$ denote the open ball, that is the set $\{(x, y) \in \mathbb{R}^2 \mid \|(x - x_o, y - y_o)\| < r\}$. Let $S_r(x_o, y_o) = \{(x, y) \in \mathbb{R}^2 \mid \|(x - x_o, y - y_o)\| = r\}$ denote the boundary of the set $D_r(x_o, y_o)$.

Now we illustrate with an example that the straightforward generalization of Rolle's theorem to higher dimensions fails. Consider the vector valued function $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by

$$f(x, y) = (x^3 + xy^2 - x, yx^2 + y^3 - y).$$

Evidently, the function f is continuous on the closed unit ball $D_1(0, 0)$ and differentiable on the open unit ball $B_1(0, 0)$. It is easy to see that $f(x, y) = 0$ for all $(x, y) \in S_1(0, 0)$, that is f is constant on the boundary of the closed unit ball. If the straightforward generalization of the mean value theorem is true, then we should be able to show that the derivative of f is zero at some point in the open unit ball $B_1(0, 0)$. Differentiating $f(x, y)$, we get

$$f'(x, y) = \begin{pmatrix} \frac{\partial(x^3 + xy^2 - x)}{\partial x} & \frac{\partial(x^3 + xy^2 - x)}{\partial y} \\ \frac{\partial(yx^2 + y^3 - y)}{\partial x} & \frac{\partial(yx^2 + y^3 - y)}{\partial y} \end{pmatrix}$$

which is

$$f'(x, y) = \begin{pmatrix} 3x^2 + y^2 - 1 & 2xy \\ 2xy & x^2 + 3y^2 - 1 \end{pmatrix}$$

and hence

$$f'(x, y) \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus the mean value theorem fails for vector valued functions on the plane.

Next, we present an important but simple result due to Furi and Martelli (1995). We need some propositions which we state without proof.

Proposition 5.1 *Let $D_r(x_o, y_o)$ be a closed disk in the plane \mathbb{R}^2 and let $f : D_r(x_o, y_o) \rightarrow \mathbb{R}$. If f has an extreme point at $(\eta, \xi) \in B_r(x_o, y_o)$ and it is differentiable at (η, ξ) , then*

$$f'(\eta, \xi) = (0, 0).$$

Proposition 5.2 *Let $D_r(x_o, y_o)$ be a closed disk in the plane \mathbb{R}^2 and let $f : D_r(x_o, y_o) \rightarrow \mathbb{R}$ be continuous. Then f has a maximum and a minimum on $D_r(x_o, y_o)$.*

Note that the above two propositions are identical to their one-dimensional version stated in Proposition 2.1 and Proposition 2.2 of Chapter 2.

For clarity and notational simplicity we shall not distinguish between column and row displays, and in the remaining portion of this section, we shall display all column vectors as row vectors.

Theorem 5.12 *Let $D_r(x_o, y_o)$ be a closed disk in the plane \mathbb{R}^2 and let the vector valued function $f : D_r(x_o, y_o) \rightarrow \mathbb{R}$ be continuous on $D_r(x_o, y_o)$ and differentiable on $B_r(x_o, y_o)$. Assume that there exists a nonzero vector (u, v) in \mathbb{R}^2 such that the inner product $\langle (u, v), f(x, y) \rangle$ is constant on the boundary of $D_r(x_o, y_o)$. Then there exists a vector $(\eta, \xi) \in B_r(x_o, y_o)$ such that*

$$\left\langle (u, v), \begin{pmatrix} \frac{\partial(f_1(\eta, \xi))}{\partial x} & \frac{\partial(f_1(\eta, \xi))}{\partial y} \\ \frac{\partial(f_2(\eta, \xi))}{\partial x} & \frac{\partial(f_2(\eta, \xi))}{\partial y} \end{pmatrix}(x, y) \right\rangle = 0$$

for all $(x, y) \in \mathbb{R}^2$.

Proof: Let (u, v) be a nonzero vector in \mathbb{R}^2 . For a temporarily fixed (u, v) , we define a function g from $D_r(x_o, y_o)$ into \mathbb{R} by

$$g(x, y) = \langle (u, v), (f_1(x, y), f_2(x, y)) \rangle.$$

Then by hypothesis, g is constant on the boundary of the disk $D_r(x_o, y_o)$. By Proposition 5.1, the function g has a maximum and a minimum in $D_r(x_o, y_o)$. Since g is constant on the boundary one of these extrema is achieved at an interior point (η, ξ) of $D_r(x_o, y_o)$. By Proposition 5.2, we get

$$g'(\eta, \xi) = (0, 0).$$

Hence, we have

$$\langle (u, v), f'(\eta, \xi)(x, y) \rangle = (0, 0)$$

that is

$$\left\langle (u, v), \begin{pmatrix} \frac{\partial(f_1(\eta, \xi))}{\partial x} & \frac{\partial(f_1(\eta, \xi))}{\partial y} \\ \frac{\partial(f_2(\eta, \xi))}{\partial x} & \frac{\partial(f_2(\eta, \xi))}{\partial y} \end{pmatrix}(x, y) \right\rangle = 0$$

for all $(x, y) \in \mathbb{R}^2$. This completes the proof.

Now we deduce Lagrange's theorem from the above theorem.

Corollary 5.1 *If $f : [a, b] \rightarrow \mathbb{R}$ is continuous on $[a, b]$ and differentiable $]a, b[$, then there exists a point $\eta \in]a, b[$ such that*

$$f(b) - f(a) = (b - a) f'(\eta).$$

Proof: Since $b - a > 0$, we have

$$[f(b) - f(a)]^2 + [b - a]^2 > 0.$$

Define $S : [a, b] \rightarrow \mathbb{R}^2$ by

$$S(t) = (t, f(t)).$$

Let

$$(u, v) = (f(b) - f(a), a - b).$$

Then

$$\langle v, S(a) \rangle = \langle v, S(b) \rangle = af(b) - bf(a).$$

That is $\langle v, S(x) \rangle$ is constant on the boundary. By the above theorem, there exists a point $\eta \in]a, b[$ such that

$$\langle v, S'(\eta) \rangle t = 0$$

for all $t \in \mathbb{R}$. This yields

$$[f(b) - f(a) + (a - b) f'(\eta)] t = 0.$$

Hence if $t \neq 0$, then

$$f(b) - f(a) = (b - a) f'(\eta).$$

The proof of the corollary is now complete.

Cauchy's mean value theorem can also be deduced from Theorem 5.12.

Corollary 5.2 *Suppose $a < b$. If $f, g : [a, b] \rightarrow \mathbb{R}$ are continuous on $[a, b]$ and differentiable $]a, b[$, then there exists a point $\eta \in]a, b[$ such that*

$$[f(b) - f(a)] g'(\eta) = [g(b) - g(a)] f'(\eta).$$

Proof: If $f(a) = f(b)$ and $g(a) = g(b)$ then there is nothing to prove.
Assume

$$[f(b) - f(a)]^2 + [g(b) - g(a)]^2 > 0.$$

Define $S : [a, b] \rightarrow \mathbb{R}^2$ by

$$S(t) = (g(t), f(t)).$$

Let

$$(u, v) = (f(b) - f(a), g(a) - g(b)).$$

Then

$$\langle u, S(a) \rangle = \langle v, S(b) \rangle = g(a) f(b) - g(b) f(a).$$

That is $\langle u, S(x) \rangle$ is constant on the boundary. By the Theorem 5.12, there exists a point $\eta \in]a, b[$ such that

$$\langle u, S'(\eta) \rangle = t = 0$$

for all $t \in \mathbb{R}$. This yields

$$[f(b) - f(a)]g'(\eta) + [g(b) - g(a)]f'(\eta) = 0.$$

Hence if $t \neq 0$, then

$$[f(b) - f(a)]g'(\eta) + [g(b) - g(a)]f'(\eta) = 0.$$

The proof of the theorem is now complete.

Theorem 5.12 can be generalized to vector valued functions from \mathbb{R}^m into \mathbb{R}^n . The proof of the following theorem is similar to the Theorem 5.12 and we leave this to the reader.

Theorem 5.13 Let $D_r(x_o)$ be a closed disk centered at $x_o \in \mathbb{R}^m$ and let the vector valued function $f : D_r(x_o) \rightarrow \mathbb{R}^n$ be continuous on $D_r(x_o)$ and differentiable on $B_r(x_o)$. Assume that there exists a nonzero vector u in \mathbb{R}^n such that the inner product $\langle u, f(x) \rangle$ is constant on the boundary of $D_r(x_o)$. Then there exists a vector $\eta \in B_r(x_o)$ such that

$$\langle u, f'(\eta)x \rangle = 0$$

for all $x \in \mathbb{R}^m$.

The following theorem is a multidimensional version of Flett's mean value theorem.

Theorem 5.14 *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be continuously differentiable in some open convex set $D \subset \mathbb{R}^n$ containing the points a and b . Suppose that there exists a vector $v \in \mathbb{R}^m$ such that $v f'(a) = v f'(b)$, then there is a point η on the interval $]a, b[$ such that*

$$\langle v, [f'(\eta)(\eta - a) - (f(\eta) - f(a))] \rangle = 0.$$

Proof: We first define an auxiliary function $\phi : [0, 1] \rightarrow \mathbb{R}$ by

$$\phi(t) = v f(a + t(b - a)).$$

Apparently, ϕ is continuously differentiable on $]0, 1[$ and

$$\phi'(t) = \langle v f'(a + t(b - a)) (b - a) \rangle.$$

Hence $\phi'(0) = \langle v, f'(a)(b - a) \rangle = \langle v, f'(b)(b - a) \rangle = \phi'(1)$. Applying Flett's mean value theorem to ϕ on the interval $[0, 1]$, we obtain

$$t_o \phi'(t_o) = \phi(t_o) - \phi(0)$$

for some $t_o \in]0, 1[$. Using the fact that $\phi(0) = v f(a)$ and letting $a - t(b - a) = \eta$ into the above equality, we obtain

$$\langle v, f'(\eta)t_o(b - a) \rangle = \langle v, f(\eta) - f(a) \rangle.$$

Since $t_o(b - a) = \eta - a$, the above equality yields the asserted result and the proof is now complete.

Note that if $m = n = 1$, then the above theorem yields the original version of Flett's mean value theorem. Even when $m = 1$ we have the same geometric meaning and the above theorem is a directional version of Flett's mean value theorem.

5.5 MVTs for Functions on the Complex Plane

Throughout this section, we will use the standard notation $z = x + iy$ for $z \in \mathbb{C}$, where $x = \text{Re}(z)$ the real part of z and $y = \text{Im}(z)$ the imaginary part of z . As in the case of vector valued functions, Rolle's theorem is

also not valid for holomorphic functions of a complex variable. To see this, consider the function

$$f(z) = e^z - 1.$$

Then f is a holomorphic function and

$$f(2k\pi i) = e^{2k\pi i} - 1 = 0$$

for $k \in \mathbb{Z}$. But the derivative of f , that is

$$f'(z) = e^z$$

has no zeros in the complex plane. Hence, Rolle's theorem is false in the complex domain without any changes.

Let us consider another example. Let $f(z) = (z^2 - 1)(z - i\sqrt{3})$. It is easy to see that the polynomial f has zeros at the vertices $z = 1, z = -1$ and $z = i\sqrt{3}$ of an isosceles triangle. If Rolle's theorem were valid, the derivative of f would be zero on each side of the triangle. But the derivative of f is

$$f'(z) = 3 \left[z - \frac{i}{\sqrt{3}} \right]^2$$

which has a zero at $z = \frac{i}{\sqrt{3}}$, a point interior to the triangle.

As the second example shows, the concept of a critical point lying between two real zeros usually is replaced in the complex plane by concept of a critical point situated in some region containing the zeros of a given function. In the second example that region is a closed triangle but a circular disk may be the most convenient choice.

The efforts made towards devising some complex plane counterparts to Rolle's theorem began about one hundred and fifty years ago with Gauss and have continued to the present day. Jean Dieudonne (1930) published a necessary and sufficient condition for the existence of a zero of $f'(z)$ in the interior of a circular disk with diameter ab when f is holomorphic and $f(a) = f(b) = 0$. Marsden (1985) furnished some results about the relative locations of the zeros of a complex polynomial and zeros of its derivative. Marsden concluded in his 1985 paper that none of the results so far has the generality and simplicity of Rolle's theorem. Hence it remains a challenge for the future to find a true analogue of Rolle's theorem in complex domain.

In 1992, Evard and Jafari (1992) gave the following generalization of Rolle's theorem to holomorphic functions of a complex variable.

Theorem 5.15 Let f be a holomorphic function defined on an open convex subset D of \mathbb{C} . Let $a, b \in D$ be such that $f(a) = f(b) = 0$ and $a \neq b$. Then there exist z_1, z_2 on the line segment $[a, b]$ such that

$$\operatorname{Re}(f'(z_1)) = 0 \quad \text{and} \quad \operatorname{Im}(f'(z_2)) = 0.$$

Proof: Let $a_1 = \operatorname{Re}(a)$, $a_2 = \operatorname{Im}(a)$, $b_1 = \operatorname{Re}(b)$, $b_2 = \operatorname{Im}(b)$ and let $u(z) = \operatorname{Re}(f(z))$ and $v(z) = \operatorname{Im}(f(z))$ for $z \in D$. We define $\phi : [0, 1] \rightarrow \mathbb{R}$ by

$$\phi(t) = (b_1 - a_1) u(a + t(b - a)) + (b_2 - a_2) v(a + t(b - a))$$

for every $t \in [0, 1]$. Since $f(a) = f(b)$, we get

$$u(a) = u(b) = v(a) = v(b) = 0.$$

Consequently, $\phi(0) = 0 = \phi(1)$. Applying Rolle's theorem to ϕ on $[0, 1]$, we obtain

$$\phi'(t_1) = 0$$

for some $t_1 \in]0, 1[$. Letting $z_1 = a + t_1(b - a)$, we get from the above equation

$$(b_1 - a_1) \left[(b_1 - a_1) \frac{\partial u(z_1)}{\partial x} + (b_2 - a_2) \frac{\partial u(z_1)}{\partial y} \right] \\ + (b_2 - a_2) \left[(b_1 - a_1) \frac{\partial v(z_1)}{\partial x} + (b_2 - a_2) \frac{\partial v(z_1)}{\partial y} \right] = 0.$$

By the Cauchy-Riemann equations, that is

$$\frac{\partial u(z)}{\partial x} = \frac{\partial v(z)}{\partial y} \quad \text{and} \quad \frac{\partial u(z)}{\partial y} = -\frac{\partial v(z)}{\partial x}$$

it follows that

$$\frac{\partial u(z_1)}{\partial x} [(b_1 - a_1)^2 + (b_2 - a_2)^2] = 0.$$

Thus

$$\frac{\partial u(z_1)}{\partial x} = 0$$

which is

$$\operatorname{Re}(f'(z_1)) = 0.$$

By applying this first part of the theorem to the function $g = -i f$ we see that there exists a $z_2 \in]a, b[$ such that

$$0 = \operatorname{Re}(g'(z_2)) = \frac{\partial v(z_2)}{\partial x} = -\frac{\partial u(z_2)}{\partial y} = \operatorname{Im}(f'(z_2)).$$

This completes the proof of the theorem.

Using the above theorem we now establish the complex mean value theorem.

Theorem 5.16 *Let f be a holomorphic function defined on an open convex subset D of \mathbb{C} . Let a and b be two distinct points in D . Then there exist $z_1, z_2 \in]a, b[$ such that*

$$\operatorname{Re}(f'(z_1)) = \operatorname{Re}\left(\frac{f(b) - f(a)}{b - a}\right) \quad \text{and} \quad \operatorname{Im}(f'(z_2)) = \operatorname{Im}\left(\frac{f(b) - f(a)}{b - a}\right).$$

Proof: Let

$$g(z) = f(z) - f(a) - \frac{f(b) - f(a)}{b - a} (z - a)$$

for every $z \in D$. Obviously $g(a) = g(b) = 0$. By the above theorem, there exist $z_1, z_2 \in]a, b[$ such that

$$\operatorname{Re}(g'(z_1)) = 0$$

and

$$\operatorname{Im}(g'(z_2)) = 0.$$

Using the definition of g , we get

$$g'(z) = f'(z) - \frac{f(b) - f(a)}{b - a}$$

for every $z \in D$. Therefore

$$0 = \operatorname{Re}(g'(z_1)) = \operatorname{Re}(f'(z_1)) - \operatorname{Re}\left(\frac{f(b) - f(a)}{b - a}\right)$$

and

$$0 = \operatorname{Im}(g'(z_2)) = \operatorname{Im}(f'(z_2)) - \operatorname{Im}\left(\frac{f(b) - f(a)}{b - a}\right).$$

The proof of the theorem is now complete.

Flett's theorem is also not valid for holomorphic functions of a complex variable. To see this, consider the function

$$f(z) = e^z - z.$$

Then f is holomorphic and its derivative is given by

$$f'(z) = e^z - 1.$$

Therefore, we have

$$f'(2k\pi i) = e^{2k\pi i} - 1 = 0$$

for $k \in \mathbb{Z}$. Hence, in particular

$$f'(2\pi i) = f'(4\pi i),$$

that is, the derivatives of f at the end points of the interval $[2\pi i, 4\pi i]$ are equal. Now we show that

$$f'(z) = \frac{f(z) - f(2\pi i)}{z - 2\pi i}$$

has no solution on the interval $]2\pi i, 4\pi i[$. Rewriting the above equation, one obtains

$$[1 + 2\pi i - z] e^z = 1,$$

which is

$$1 - i(y - 2\pi) = \cos y - i \sin y.$$

Comparing the real and imaginary parts, we get the following system of equations

$$\cos y = 1 \quad \text{and} \quad \sin y = y - 2\pi.$$

It is easy to check that this system of equations has no solution on the interval $]2\pi, 4\pi[$ and consequently there is no complex number on the open

interval $]2\pi i, 4\pi i[$ such that $[1+2\pi i - z] e^z = 1$ has a solution. Thus Flett's theorem is false in the complex domain without any changes.

For distinct a and b in \mathbb{C} , $[a, b]$ denotes the set $\{a + t(b - a) \mid t \in [0, 1]\}$ and it is referred to as a *line segment* or a *closed interval* in \mathbb{C} . Similarly, by $]a, b[$ we mean the set $\{a + t(b - a) \mid t \in]0, 1[\}$. Now we give a generalization of Flett's theorem for holomorphic functions of a complex variable in the spirit of Evard and Jafari (1992). This theorem is due to Davitt, Powers, Riedel and Sahoo (1998).

Theorem 5.17 . *Let f be a holomorphic function defined on an open convex subset D of \mathbb{C} . Let a and b be two distinct points in D . Then there exist $z_1, z_2 \in]a, b[$ such that*

$$\operatorname{Re}(f'(z_1)) = \frac{\langle b - a, f(z_1) - f(a) \rangle}{\langle b - a, z_1 - a \rangle} + \frac{1}{2} \frac{\operatorname{Re}(f'(b) - f'(a))}{b - a} (z_1 - a)$$

and

$$\operatorname{Im}(f'(z_2)) = \frac{\langle b - a, -i[f(z_2) - f(a)] \rangle}{\langle b - a, z_2 - a \rangle} + \frac{1}{2} \frac{\operatorname{Im}(f'(b) - f'(a))}{b - a} (z_2 - a),$$

where $\langle \cdot, \cdot \rangle : \mathbb{C}^2 \rightarrow \mathbb{R}$ denotes

$$\langle u, v \rangle = \operatorname{Re}(u) \operatorname{Re}(v) + \operatorname{Im}(u) \operatorname{Im}(v).$$

Proof: Let $a_1 = \operatorname{Re}(a)$, $a_2 = \operatorname{Im}(a)$, $b_1 = \operatorname{Re}(b)$, $b_2 = \operatorname{Im}(b)$, $u(z) = \operatorname{Re}(f(z))$ and $v(z) = \operatorname{Im}(f(z))$ for $z \in D$. We define $\phi : [0, 1] \rightarrow \mathbb{R}$ by

$$\phi(t) = \langle b - a, f(a + t(b - a)) \rangle \tag{5.23}$$

which is

$$\phi(t) = (b_1 - a_1) u(a + t(b - a)) + (b_2 - a_2) v(a + t(b - a))$$

for every $t \in [0, 1]$. Taking the derivative of ϕ , we have

$$\phi'(t) = (b_1 - a_1) \left[(b_1 - a_1) \frac{\partial u(z)}{\partial x} + (b_2 - a_2) \frac{\partial u(z)}{\partial y} \right]$$

$$+ (b_2 - a_2) \left[(b_1 - a_1) \frac{\partial v(z)}{\partial x} + (b_2 - a_2) \frac{\partial v(z)}{\partial y} \right],$$

where $z = a + t(b - a)$. Since f is holomorphic, f satisfies the Cauchy-Riemann equations, that is

$$\frac{\partial u(z)}{\partial x} = \frac{\partial v(z)}{\partial y} \quad \text{and} \quad \frac{\partial u(z)}{\partial y} = -\frac{\partial v(z)}{\partial x}.$$

Thus, we obtain

$$\phi'(t) = \frac{\partial u(z)}{\partial x} [(b_1 - a_1)^2 + (b_2 - a_2)^2] = |b - a|^2 \operatorname{Re}(f'(z)). \quad (5.24)$$

Applying Theorem 5.2 to ϕ on $[0, 1]$, we obtain

$$t_1 \phi'(t_1) = \phi(t_1) - \phi(0) + \frac{1}{2} \frac{\phi'(1) - \phi'(0)}{1 - 0} (t_1 - 0)^2$$

for some $t_1 \in]0, 1[$. Thus

$$t_1 |b - a|^2 \operatorname{Re}(f'(z_1)) = \phi(t_1) - \phi(0) + \frac{1}{2} [\phi'(1) - \phi'(0)] t_1^2,$$

where $z_1 = a + t_1(b - a)$. Further, since $z_1 = a + t_1(b - a)$ and $t_1 \in [0, 1]$, we have

$$\begin{aligned} t_1 |b - a|^2 &= t_1 (b_1 - a_1)^2 + t_1 (b_2 - a_2)^2 \\ &= t_1 (b_1 - a_1) \operatorname{Re}(b - a) + t_1 (b_2 - a_2) \operatorname{Im}(b - a) \\ &= \operatorname{Re}(z_1 - a) \operatorname{Re}(b - a) + \operatorname{Im}(z_1 - a) \operatorname{Im}(b - a) \\ &= \langle b - a, z_1 - a \rangle. \end{aligned}$$

Hence the equation $t_1 |b - a|^2 \operatorname{Re}(f'(z_1)) = \phi(t_1) - \phi(0) + \frac{1}{2} [\phi'(1) - \phi'(0)] t_1^2$ reduces to

$$\operatorname{Re}(f'(z_1)) = \frac{\phi(t_1) - \phi(0)}{t_1 |b - a|^2} + \frac{1}{2} \frac{\phi'(1) - \phi'(0)}{|b - a|^2} t_1.$$

Using (5.23), (5.24) and the fact that $z_1 = a + t_1(b - a)$ in the above equation, we obtain

$$\operatorname{Re}(f'(z_1)) = \frac{\langle b - a, f(z_1) - f(a) \rangle}{\langle b - a, z_1 - a \rangle} + \frac{1}{2} \frac{\operatorname{Re}(f'(b) - f'(a))}{b - a} (z_1 - a).$$

Letting $g = -if$, we have

$$\operatorname{Re}(g'(z)) = \frac{\partial v(z)}{\partial x} = -\frac{\partial u(z)}{\partial y} = \operatorname{Im}(f'(z)). \quad (5.25)$$

Now applying the first part to g , we obtain

$$\operatorname{Re}(g'(z_2)) = \frac{\langle b - a, g(z_2) - g(a) \rangle}{\langle b - a, z_2 - a \rangle} + \frac{1}{2} \frac{\operatorname{Re}(g'(b) - g'(a))}{b - a} (z_2 - a)$$

for some $z_2 \in]a, b[$. By (5.25) the above yields

$$\operatorname{Im}(f'(z_2)) = \frac{\langle b - a, -i[f(z_2) - f(a)] \rangle}{\langle b - a, z_2 - a \rangle} + \frac{1}{2} \frac{\operatorname{Im}(f'(b) - f'(a))}{b - a} (z_2 - a)$$

and the proof is now complete.

The following corollary is transparent from the above theorem and it is the complex version of Flett's mean value theorem.

Corollary 5.3 . *Let f be a holomorphic function defined on an open convex subset D of \mathbb{C} . Let a and b be two distinct points in D , and $f'(a) = f'(b)$. Then there exist $z_1, z_2 \in]a, b[$ such that*

$$\operatorname{Re}(f'(z_1)) = \frac{\langle b - a, f(z_1) - f(a) \rangle}{\langle b - a, z_1 - a \rangle}$$

and

$$\operatorname{Im}(f'(z_2)) = \frac{\langle b - a, -i[f(z_2) - f(a)] \rangle}{\langle b - a, z_2 - a \rangle},$$

where $\langle \cdot, \cdot \rangle : \mathbb{C}^2 \rightarrow \mathbb{R}$ denotes

$$\langle u, v \rangle = \operatorname{Re}(u) \operatorname{Re}(v) + \operatorname{Im}(u) \operatorname{Im}(v).$$

Returning to our previous example,

$$f(z) = e^z - z,$$

we will now find numerical approximations for z_1 and z_2 guaranteed to exist by Theorem 5.17. Let $a = 0, b = 2\pi i$. Since z_1 and z_2 have to lie on the line segment joining a and b it follows that $z_1 = iy_1$ and $z_2 = iy_2$. Further, since $\operatorname{Re}(b - a) = 0, \operatorname{Im}(b - a) = 2\pi$ and $f'(a) = f'(b)$, the first part of Theorem 5.17 yields

$$\operatorname{Re}(e^{iy_1} - 1) = \frac{2\pi \operatorname{Im}(e^{iy_1} - iy_1 - 1)}{2\pi y_1}.$$

So

$$\cos(y_1) - 1 = \frac{\sin(y_1) - y_1}{y_1}.$$

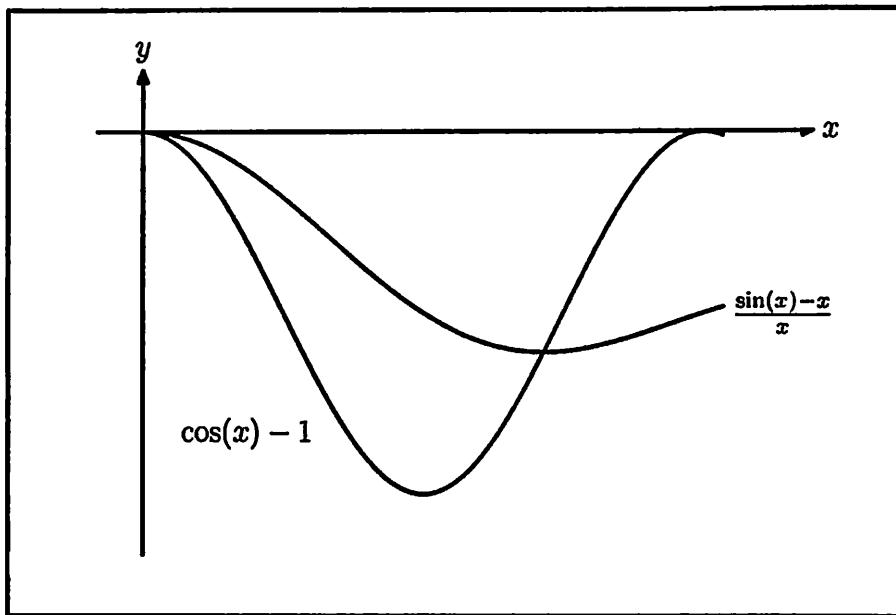


Fig. 5.3 Graph of the Functions $\cos(y_1) - 1$ and $\frac{\sin(y_1)-y_1}{y_1}$

The above nonlinear equation has a solution for $y_1 \approx 4.49341$, and by looking at the graph (see Figure 5.3), we see that it has no other solution in $]0, 2\pi[$. Thus $z_1 \approx 4.49341i$. Similarly we obtain

$$\operatorname{Im}(e^{iy_2} - 1) = -\frac{2\pi \operatorname{Re}(e^{iy_2} - iy_2 - 1)}{2\pi y_2}$$

which is

$$-\sin(y_2) = \frac{\cos(y_2) - 1}{y_2}.$$

This equation has exactly one solution in $]0, 2\pi[$, and $y_2 \approx 2.33112$ (see Figure 5.4). Thus $z_2 \approx 2.33112i$. Trahan (1966) gave a new proof of Flett's theorem replacing the boundary condition $f'(a) = f'(b)$ by

$$[f'(a) - m][f'(b) - m] > 0,$$

where $m = \frac{f(b) - f(a)}{b - a}$. Similarly, the same conclusion of the Corollary 5.3 can be achieved by replacing the boundary condition $f'(a) = f'(b)$ by

$$[\operatorname{Re}(f'(a)) - m_1][\operatorname{Re}(f'(b)) - m_1] > 0$$

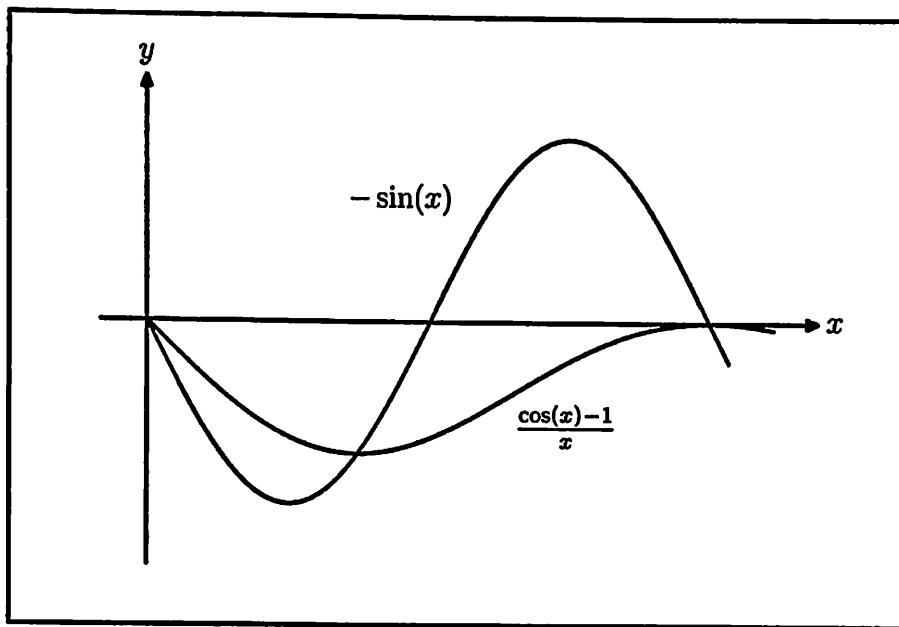


Fig. 5.4 Graph of the Functions $-\sin(y_2)$ and $\frac{\cos(y_2)-1}{y_2}$

and

$$[Im(f'(a)) - m_2][Im(f'(b)) - m_2] > 0,$$

where $m_1 = \frac{\langle b-a, f(b)-f(a) \rangle}{\langle b-a, b-a \rangle}$ and $m_2 = \frac{\langle b-a, -i[f(b)-f(a)] \rangle}{\langle b-a, b-a \rangle}$. The proof follows by defining

$$\phi(t) = \langle b-a, f(a+t(b-a)) \rangle$$

as in the Theorem 5.17 and applying Trahan's result to ϕ . We leave the details to the reader.

In the following theorem, we present a local interpretation of Lagrange's mean value theorem in the complex plane. However, the mean value z lies near the points a and b but not necessarily on the line segment joining them. The proof of the following local version of Lagrange's mean value theorem for holomorphic functions is due to Samuelsson (1973) and it uses Rouche's theorem.

Theorem 5.18 *Let f be analytic in a region $G \subset \mathbb{C}$ containing a . Then there is a neighborhood $N(a)$ such that for every $b \in N(a)$ there is a z in*

the open disk D about $\frac{b-a}{2}$ of radius $\frac{1}{2}|b-a|$ with

$$f(b) - f(a) = f'(z)(b-a). \quad (5.26)$$

Proof: We first note that if f is linear then equation (5.26) holds for any $z \in G$. Now assume that f is not linear, using Taylor's theorem, we have that there is an integer $k \geq 1$ and a function $h(z)$ analytic in G with $h(a) \neq 0$ such that

$$f(z) = f(a) + (z-a)f'(a) + (b-a)^{k+1}h(z) \quad (5.27)$$

for all $z \in G$. Now we choose a neighborhood N_1 of a inside of G and let M be the maximum of $|h'(z)|$ on the closure of N_1 . Replacing f by $\frac{f}{M}$ and h by $\frac{h}{M}$ and labeling them again as f and h respectively, we can assume that $|h'(z)| \leq 1$ throughout the domain of interest. Similarly, by choosing z close enough to a , we may assume that $\frac{1}{2}|h(a)| \leq |h(z)|$.

To prove the theorem, it suffices to show that

$$f'(z) - \frac{f(b) - f(a)}{b-a} \quad (5.28)$$

has a zero in the open disk D about $\frac{b-a}{2}$ of radius $\frac{1}{2}|b-a|$. We proceed by writing the function in equation (5.28) as a sum of two functions, one of which has a known zero in this disk and dominates the other. Then we will use Rouché's theorem to finish the proof.

Using equation (5.27) we can calculate the derivative of f to be

$$f'(z) = f'(a) + (k-1)(z-a)^k h(z) + (z-a)^{k+1}h'(z), \quad (5.29)$$

and further we also get

$$f(b) = f(a) + (b-a)f'(a) + (b-a)^{k+1}h(b). \quad (5.30)$$

Using (5.30) we get

$$\frac{f(b) - f(a)}{b-a} = f'(a) + (b-a)h(b), \quad (5.31)$$

and subtracting this from (5.29) yields

$$f'(b) - \frac{f(b) - f(a)}{b-a} = \phi(z) + h(b)\psi(z), \quad (5.32)$$

where

$$\phi(z) = (z - a)^{k+1} h'(z) + (k+1)(z - a)^k (h(z) - h(a)) \quad (5.33)$$

$$\psi(z) = (k+1)(z - a)^k - (b - a)^k, \quad (5.34)$$

we will now show that ϕ and ψ have the same number of zeros in a suitably restricted area. If b is in $N_1 = \left\{ z \mid |z - a| < \frac{|h(a)|}{2} \frac{1}{k+2} \right\}$, then we note that ψ , being a k -th degree polynomial has exactly one zero in the sector D_s of D , where

$$D_s = \left\{ z \mid \left| z - \frac{1}{2}(z_1 - a) \right| < \frac{1}{2}|b - a|; \text{ with } \arg \left(\frac{z - a}{b - a} \right) < \frac{\pi}{k} \right\}.$$

If z is on the boundary ∂D_s , then we have

$$\begin{aligned} |\phi(z)| &\leq |z - a|^{k+1} |h'(z)| + (k+1)|z - a|^k \left| \int_b^z h'(\xi) d\xi \right| \\ &\leq (k+2)|b - a|^{k+1}. \end{aligned} \quad (5.35)$$

The latter inequality holds since $|h'(z)| \leq 1$ and $|z - a| \leq |b - a|$. To estimate ψ we need to split ∂D_s . If $z = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)e^{i2\theta}$, for $|\theta| \leq \frac{\pi}{k}$, (that is z is on the circular arc), then, using $e^{i2\theta} = \cos(2\theta) + i\sin(2\theta)$ we get

$$\frac{|\psi(z)|^2}{|b - a|^{2k}} = 1 + (k+1)((k+1)\cos^k(\theta) - 2\cos(k\theta))\cos^k(\theta). \quad (5.36)$$

Since $(k+1)\cos^k(\theta) - 2\cos(k\theta) \geq 0$ for $|\theta| \leq \frac{\pi}{k}$, we obtain

$$\frac{|\psi(z)|^2}{|b - a|^{2k}} \geq 1 \quad \text{or} \quad |\psi(z)| \geq |b - a|^k. \quad (5.37)$$

If, on the other hand z is on one of the line segments of D_s , that is $z = a + t(b - a)e^{\pm i\pi/k}$, for $0 \leq t \leq \cos(\frac{\pi}{k})$, then we obtain

$$|\psi(z)| = (1 + (k+1)t^k) |b - a|^k \geq |b - a|^k, \quad (5.38)$$

which is the same estimate as (5.37). Finally, if $b \in N_1$ and $z \in \partial D_s$ then we get

$$\left| \frac{\phi(z)}{h(b)\psi(z)} \right| \leq \frac{k+2}{|h(b)|} |b - a| < \frac{|h(a)|}{2|h(b)|} \leq 1, \quad (5.39)$$

where the last inequality holds since we assume $z \in N_1$. Thus by Rouche's theorem,

$$h(b)\psi(z) \text{ and } \phi(z) + h(b)\psi(z) = f'(z) - \frac{f(b) - f(a)}{b - a} \quad (5.40)$$

have the same number of zeros in the interior of D , namely one. This concludes the proof.

5.6 A Conjecture of Furi and Martelli

Theorem 5.13 in section 4 of this chapter due to Furi and Martelli (1995). It is valid if one replaces the range space \mathbb{R}^n by a Hilbert space \mathbb{H} . The proof is identical to the proof of Theorem 5.13. It is also true when the range space \mathbb{R}^n is replaced by a Banach space \mathbb{F} and the inner product $\langle u, f'(\eta)x \rangle$ is replaced by $\phi(f'(\eta)x)$, where ϕ is a continuous linear functional.

Furi and Martelli (1995) have conjectured that Theorem 5.13 is false if the domain of f , \mathbb{R}^m , is replaced by an infinite dimensional Banach space \mathbb{E} . Notice that the multidimensional versions of Proposition 5.2 is no longer true in \mathbb{E} . It is well known that the unit closed ball $D_1(0)$ is not compact in \mathbb{E} and consequently, there are continuous functions $f : D_1(0) \rightarrow \mathbb{R}$ such that the image of f is an open interval.

Recently Ferrer (1996) provided an example which showed that Theorem 5.13 fails in infinite dimension. This proves Furi and Martelli's conjecture to be true. To establish the conjecture he gave an example of a function f from the space l_2 of square summable real sequences into \mathbb{R} with the following properties:

- $f(x) = 0$ for all x on the unit sphere in l_2
- f is continuous on the closed unit ball in l_2
- f is differentiable on the open unit ball in l_2 .

Ferrer's proof makes use of the fact that the unit ball in the Hilbert space l_2 (with the l_2 -norm) is not compact, and thus f is constructed to be unbounded on this set. Ferrer then shows that there is no point η in the open unit ball such that $f'(\eta) = 0$.

The proof of this theorem goes well beyond the scope of this book and the interested reader is referred to the original article.

Chapter 6

Mean Value Theorems for Some Generalized Derivatives

In previous chapters, we examined various mean value theorems for functions which are differentiable in the classical sense. In many classical optimization problems one uses these derivatives. By “classical” we mean the functions involved in the problem are differentiable or continuously differentiable. Besides treating the classical smooth problems, mathematicians have gotten many impulses in the last decades from other disciplines such as economics and engineering in order to treat nonsmooth, nondifferentiable optimization problems. In many nonsmooth optimization problems generalized derivatives such as Dini derivatives and the corresponding mean value theorems are used. In this chapter, we shall study some mean value theorems and their generalizations when a function has a generalized derivative such as a Dini derivative or a symmetric derivative. The first four sections of this chapter deal with symmetric differentiation and generalization of the mean value theorem for symmetrically differentiable functions whereas the last two sections focus on Dini derivatives and various mean value theorems for these derivatives. We begin this chapter with an elementary introduction to symmetric differentiation.

6.1 Symmetric Differentiation of Real Functions

Recall that the derivative of a real function f , that is a function from the real line \mathbb{R} into itself, at a point x is given by the limit

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

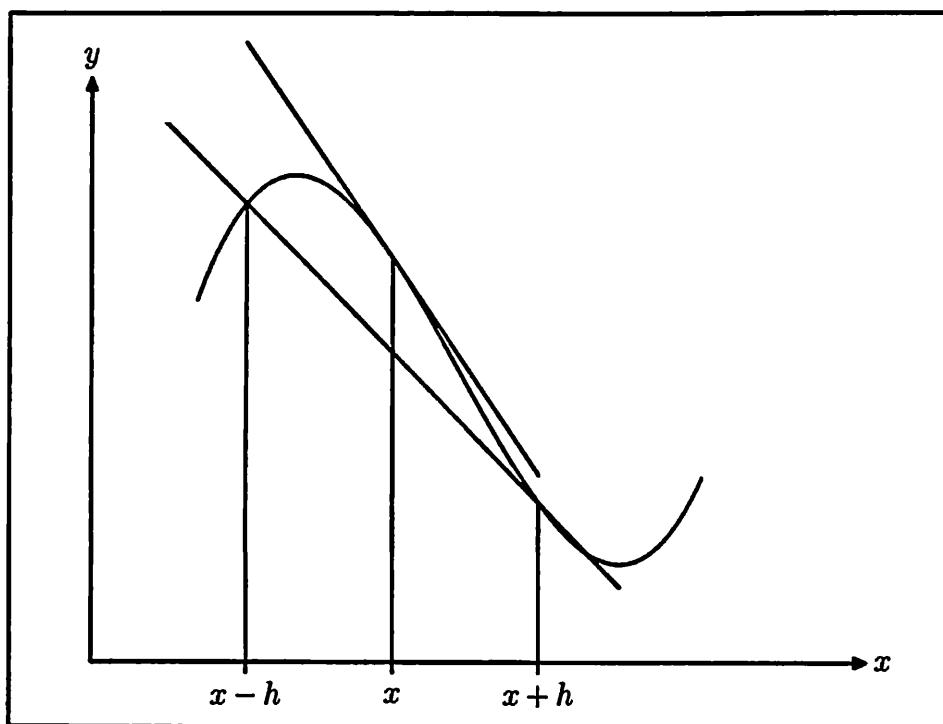


Fig. 6.1 A Geometric Interpretation of the Symmetric Derivative.

provided this limit exists. *The limit exists will mean it exists and is finite.* If a function has a finite derivative, then its graph is continuous and devoid of sharp corners and vertical tangents. In this sense, differentiable functions are functions with nice graphs.

Over the years, mathematicians have proposed many generalizations of derivatives by altering the right side of the definition of derivative stated in the previous page. If we replace the difference quotient $\frac{f(x+h)-f(x)}{h}$ by a central difference quotient $\frac{f(x+h)-f(x-h)}{2h}$, then we obtain a generalized derivative. This generalized derivative is called the symmetric derivative of the function f . Interested readers are referred to the book by Thomson (1994) for treatments on symmetric derivatives of real functions and related topics. Formally, we define it as follow.

Definition 6.1 A real function f on an interval $]a, b[$ is said to be symmetrically differentiable at a point x in $]a, b[$ if the limit

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x-h)}{2h}$$

exists.

We shall denote this limit as $f^s(x)$. If a function is *symmetrically differentiable* at every point of an interval, then we say it is symmetrically differentiable on that interval. The Figure 6.1 geometrically illustrates the notion of symmetric differentiability.

One can easily show that if a function is differentiable, then it is also symmetrically differentiable.

Theorem 6.1 *Every differentiable function is symmetrically differentiable.*

Proof: Let x be an arbitrary point. We would like to show that the symmetric derivative $f^s(x)$ exists. Since

$$\begin{aligned} f^s(x) &= \lim_{h \rightarrow 0} \frac{f(x+h) - f(x-h)}{2h} \\ &= \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{2h} + \lim_{h \rightarrow 0} \frac{f(x) - f(x-h)}{2h} \\ &= \frac{1}{2} f'(x) + \frac{1}{2} f'(x) = f'(x), \end{aligned}$$

the symmetric derivative of f , that is $f^s(x)$ exists at every point x . Hence f is symmetrically differentiable and the proof is now complete.

This theorem says that ordinary differentiability implies symmetric differentiability. However, the converse is not true as seen by the following example.

Example 6.1 The function $f(x) = |x|$ is symmetrically differentiable at zero but it is not differentiable at zero.

To see this we evaluate the limits

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \quad \text{and} \quad \lim_{h \rightarrow 0} \frac{f(x+h) - f(x-h)}{2h}$$

at the point $x = 0$. Evaluating the first limit, we get

$$\begin{aligned} f'(0) &= \lim_{h \rightarrow 0} \frac{f(0+h) - f(0)}{h} \\ &= \lim_{h \rightarrow 0} \frac{|h| - |0|}{h} \\ &= \lim_{h \rightarrow 0} \begin{cases} \frac{h}{h} & \text{if } h \geq 0 \\ \frac{-h}{h} & \text{if } h < 0 \end{cases} \\ &= \begin{cases} 1 & \text{if } h \geq 0 \\ -1 & \text{if } h < 0. \end{cases} \end{aligned}$$

Hence the limit does not exist at 0. Therefore f is not differentiable at 0.

Now we examine the symmetric differentiability of f . For this, we compute

$$\begin{aligned} f^s(0) &= \lim_{h \rightarrow 0} \frac{f(0+h) - f(0-h)}{2h} \\ &= \lim_{h \rightarrow 0} \frac{|h| - |-h|}{2h} \\ &= \lim_{h \rightarrow 0} 0 \\ &= 0. \end{aligned}$$

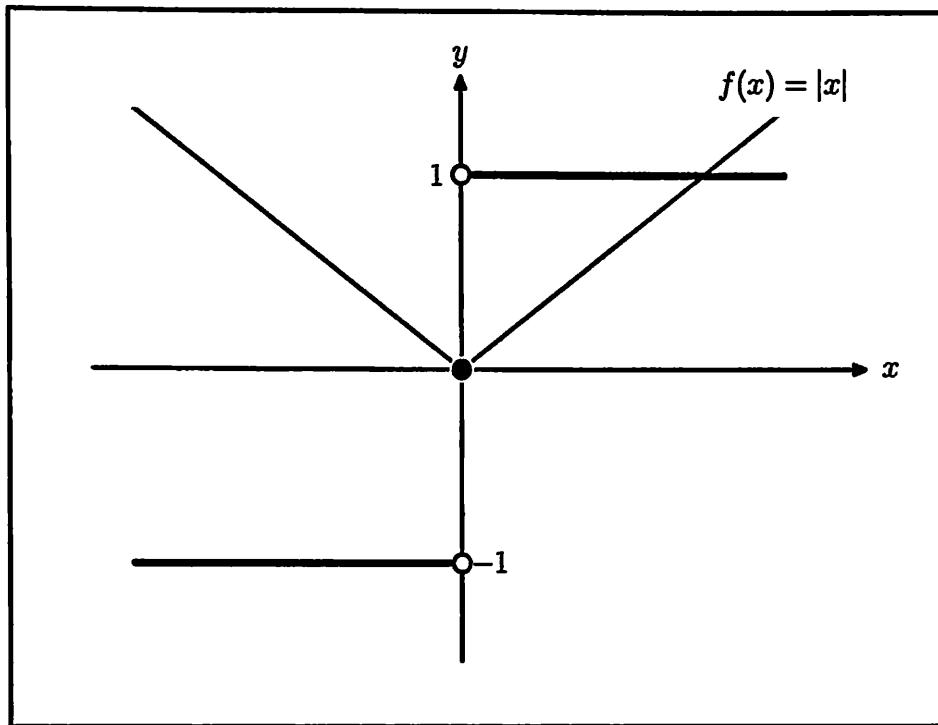
Hence the limit exists and is equal to 0. Therefore f is symmetrically differentiable at 0. The symmetric derivative of $f(x) = |x|$ is

$$f^s(x) = \begin{cases} \frac{|x|}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

The graph of $f(x)$ and $f^s(x)$ is shown in Figure 6.2.

Note that the absolute value function $f(x) = |x|$ has a sharp corner at 0 and thus it is not differentiable. However, sharp corner has no bearing on the symmetrically differentiable function. We know from our knowledge of ordinary differentiation that every differentiable function is continuous. Now we wonder if the same is true for symmetrically differentiable functions. The following example will shed some light on our query.

Example 6.2 The function $f(x) = \frac{1}{x^2}$ is symmetrically differentiable at zero but it is not defined at zero.

Fig. 6.2 Graph of $|x|$ and its Symmetric Derivative.

To see this, we evaluate the limit

$$\begin{aligned} f^s(0) &= \lim_{h \rightarrow 0} \frac{f(0+h) - f(0-h)}{2h} \\ &= \lim_{h \rightarrow 0} \frac{\frac{1}{h^2} - \frac{1}{h^2}}{2h} \\ &= \lim_{h \rightarrow 0} 0 \\ &= 0. \end{aligned}$$

Hence $\frac{1}{x^2}$ is symmetrically differentiable at $x = 0$ although it is not defined at zero. The symmetric derivative is given by

$$f^s(x) = \begin{cases} -\frac{2}{x^3} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

The function $f(x) = \frac{1}{x^2}$ is not continuous at $x = 0$. This example illustrates that, contrary to the notion of ordinary differentiability, a discontinuous function can have a symmetric derivative.

Note that by generalizing the definition of differentiability to symmetric differentiability we lost some aesthetic properties like continuity and smoothness of curves. In the next examples we illustrate that if a function has a finite jump discontinuity at a point, then it may or may not be symmetrically differentiable.

Example 6.3 The function

$$f(x) = \begin{cases} 2 & \text{if } x \neq 0 \\ 5 & \text{if } x = 0. \end{cases}$$

is symmetrically differentiable at zero but it is not continuous.

It is easy to see that f is discontinuous at 0. Since

$$f^s(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x-h)}{2h} = \lim_{h \rightarrow 0} \frac{2-2}{2h} = 0,$$

the function f symmetrically differentiable.

This example shows that if the limit of f exists as x approaches 0, then f is symmetrically differentiable at 0 even though f is discontinuous at 0. In the following example, we shall see that if a function has a finite jump discontinuity with no limit at 0, then it has no symmetric derivative.

Example 6.4 The function

$$f(x) = \begin{cases} \frac{|x|}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

is not symmetrically differentiable at zero.

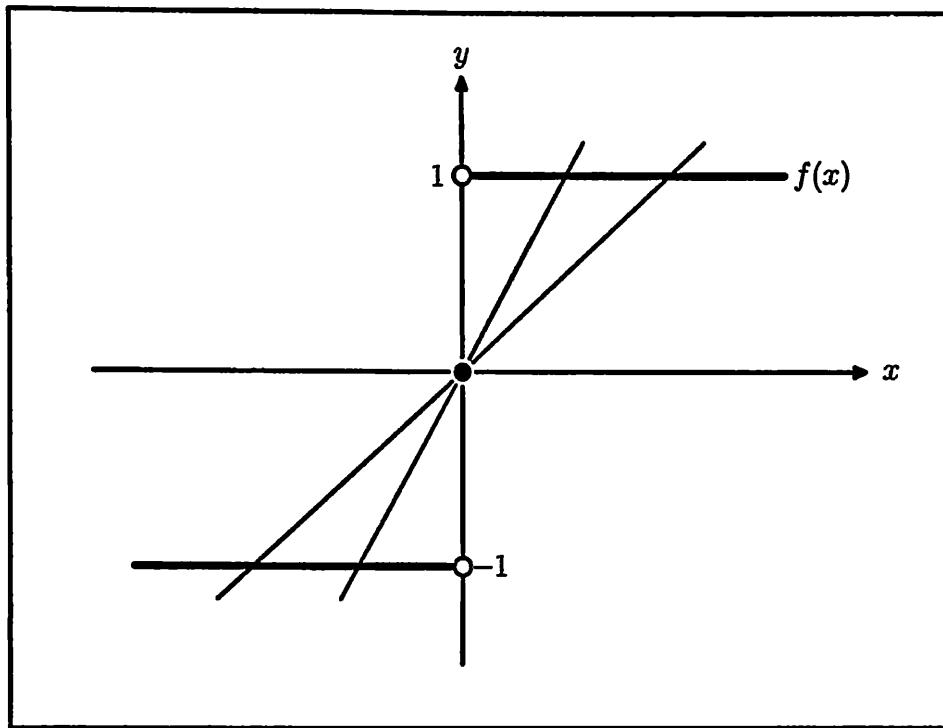


Fig. 6.3 An Example of a Function Devoid of Symmetric Derivative.

This can be shown by computing the limit

$$\begin{aligned}
 f^s(0) &= \lim_{h \rightarrow 0} \frac{f(0+h) - f(0-h)}{2h} \\
 &= \lim_{h \rightarrow 0} \frac{\frac{|h|}{h} - \frac{|-h|}{-h}}{2h} \\
 &= \lim_{h \rightarrow 0} \frac{|h|}{h^2} \\
 &= \lim_{h \rightarrow 0} \begin{cases} \frac{h}{h^2} & \text{if } h \geq 0 \\ \frac{-h}{h^2} & \text{if } h < 0 \end{cases} \\
 &= \lim_{h \rightarrow 0} \begin{cases} \frac{1}{h} & \text{if } h \geq 0 \\ -\frac{1}{h} & \text{if } h < 0. \end{cases} \\
 &= \infty.
 \end{aligned}$$

Hence this function has no symmetric derivative. The graph of this function is shown in the Figure 6.3.

6.2 A Quasi-Mean Value Theorem

In this section, we establish a quasi-mean value theorem for functions with symmetric derivatives. We will further show that every continuous function whose symmetric derivative has the Darboux property obeys the ordinary mean value theorem of Lagrange.

The ordinary mean value theorem is not true for symmetric derivatives as illustrated in the following example.

Example 6.5 The function $f(x) = |x|$ does not satisfy the ordinary mean value theorem on the interval $[-1, 2]$.

The symmetric derivative of the function $f(x) = |x|$ is given by

$$f^s(x) = \begin{cases} \frac{|x|}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Note that the range of $f^s(x)$ is the set $\{0, -1, 1\}$. The slope of the secant line is $\frac{f(2)-f(-1)}{3}$ which is $\frac{1}{3}$. Since the range of $f^s(x)$ does not contain this value therefore there is no η such that

$$f^s(\eta) = \frac{1}{3}.$$

The following lemma due to Aull (1967) is instrumental in proving the so called quasi-mean value theorem.

Lemma 6.1 Let f be continuous on the interval $[a, b]$ and let f be symmetrically differentiable on $]a, b[$. If $f(b) > f(a)$, then there exists a point $\eta \in]a, b[$ such that

$$f^s(\eta) \geq 0.$$

Further, if $f(b) < f(a)$, then there exists a point $\xi \in]a, b[$ such that

$$f^s(\xi) \leq 0.$$

Proof: Suppose $f(b) > f(a)$. Let k be a real number such that $f(a) < k < f(b)$. The set

$$\{x \in [a, b] \mid f(x) > k\}$$

is bounded from below by a . Since it is a subset of \mathbb{R} it has a greatest lower bound, say, η . Since f is continuous and k satisfies $f(a) < k < f(b)$,

therefore η is different from a and b . Let $\eta - h, \eta + h$ be an arbitrary neighborhood of η in $[a, b]$. Since η is the greatest lower bound of the set $\{x \in [a, b] \mid f(x) > k\}$, there are points in $\eta - h, \eta + h$ such that

$$f(x + h) > k$$

and

$$f(x - h) \leq k.$$

Therefore

$$f^s(\eta) = \lim_{h \rightarrow 0} \frac{f(\eta + h) - f(\eta - h)}{2h} \geq 0.$$

Similarly, it can be shown that if $f(a) > f(b)$, then there exists $\xi \in]a, b[$ such that

$$f^s(\xi) \leq 0.$$

The proof of this lemma is now complete.

The following theorem of Aull (1967) can be considered as a version of Rolle's theorem for symmetrically differentiable functions.

Theorem 6.2 *Let f be continuous on $[a, b]$ and symmetrically differentiable on $]a, b[$. Suppose $f(a) = f(b) = 0$. Then there exist η and ξ in $]a, b[$ such that*

$$f^s(\eta) \geq 0$$

and

$$f^s(\xi) \leq 0.$$

Proof: If $f \equiv 0$, then the theorem is obviously true. Hence, we assume that $f \not\equiv 0$. Since f is continuous and $f(a) = f(b) = 0$, there are points x_1 and x_2 such that

$$f(x_1) > 0 \quad \text{and} \quad f(x_2) < 0 \tag{6.1}$$

or

$$f(x_1) < 0 \quad \text{and} \quad f(x_2) > 0 \tag{6.2}$$

or

$$f(x_1) > 0 \quad \text{and} \quad f(x_2) > 0 \quad (6.3)$$

or

$$f(x_1) < 0 \quad \text{and} \quad f(x_2) < 0. \quad (6.4)$$

If the inequalities in (6.1) are true, then we apply Lemma 6.1 to f on the interval $[a, x_1]$ to obtain

$$f^s(\eta) \geq 0$$

for some $\eta \in]a, x_1[\subset]a, b[$. Again applying Lemma 6.1 to f on the interval $[a, x_2]$, we obtain

$$f^s(\xi) \leq 0$$

for some $\xi \in]a, x_2[\subset]a, b[$. The other cases can be handled in a similar manner and the proof of the theorem is now complete.

Now we prove the quasi-mean value theorem for symmetrically differentiable function.

Theorem 6.3 *Let f be continuous on $[a, b]$ and symmetrically differentiable on $]a, b[$. Then there exist η and ξ in $]a, b[$ such that*

$$f^s(\eta) \leq \frac{f(b) - f(a)}{b - a} \leq f^s(\xi).$$

Proof: Define g by

$$g(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Then $g(a) = g(b) = 0$. Applying the above theorem to g on the interval $[a, b]$, we get

$$g^s(\eta) \leq 0 \quad \text{and} \quad g^s(\xi) \geq 0. \quad (6.5)$$

From (6.5) and the definition of g , we obtain

$$f^s(\eta) \leq \frac{f(b) - f(a)}{b - a} \leq f^s(\xi)$$

and the proof of the theorem is now complete.

From the above theorem, we see that the slope of the secant line through the points $(a, f(a))$ and $(b, f(b))$ can not be equal to the value of the symmetric derivative of f at an intermediate point. Now the question arises what additional condition or conditions should be imposed on the symmetric derivative of f so that the regular mean value theorem will hold for functions that are symmetrically differentiable. We will show that if the symmetric derivative of f has the Darboux property then the regular mean value theorem will hold.

Definition 6.2 A real valued function f defined on the interval $[a, b]$ is said to have the *Darboux property* if whenever η and ξ are in $[a, b]$, and y is any number between $f(\eta)$ and $f(\xi)$, then there exists a number γ between η and ξ such that $y = f(\gamma)$.

We know from calculus that every continuous function has the intermediate value property, that is the Darboux property. The Darboux property was believed by some mathematicians in 19th century to be equivalent to the property of continuity. In 1875, Darboux showed that this belief was not justified. It can be shown that a function having the Darboux property can be discontinuous everywhere.

Theorem 6.4 Let f be continuous on $[a, b]$ and symmetrically differentiable on $]a, b[$. If the symmetric derivative of f has the Darboux property, then there exists γ in $]a, b[$ such that

$$f^s(\gamma) = \frac{f(b) - f(a)}{b - a}.$$

Proof: By the above theorem, we obtain η and ξ in $]a, b[$ such that

$$f^s(\eta) \leq \frac{f(b) - f(a)}{b - a} \leq f^s(\xi).$$

Since $f^s(x)$ has the Darboux property, there exists a γ in $]a, b[$ such that

$$f^s(\gamma) = \frac{f(b) - f(a)}{b - a}.$$

This completes the proof.

6.3 An Application

Since symmetrically differentiable functions are not necessarily differentiable, the question arises what additional conditions should be imposed on the function to make it differentiable. We have seen that continuity of the function along with symmetric differentiability does not imply differentiability. In this section, we show by means of the quasi-mean value theorem that if $f(x)$ and $f^s(x)$ are both continuous, then f is differentiable. The following result is due to Aull (1967).

Theorem 6.5 *Let $f(x)$ be continuous and symmetrically differentiable on $]a, b[$. If the symmetric derivative of f is continuous on $]a, b[$, then $f'(x)$ exists and*

$$f'(x) = f^s(x).$$

Proof: Choose h to be sufficiently small so that $a < x + h < b$. Since f^s is continuous, it has the Darboux property. Applying the mean value theorem to f on $[x, x + h]$, we have

$$f^s(\eta) = \frac{f(x + h) - f(x)}{h}$$

for some $\eta \in]x, x + h[$. Taking limit of both sides as $h \rightarrow 0$ and knowing that the limit of the left side exists, one obtains

$$f^s(x) = f'(x).$$

This completes the proof.

This above theorem can be made even stronger. We leave the proof of the following theorem to the reader.

Theorem 6.6 *Let $f^s(x)$ be continuous at a point $x = a$ and let $f(x)$ be continuous in a neighborhood of a . Then $f'(a)$ exists and*

$$f'(a) = f^s(a).$$

This shows that the continuity of $f^s(x)$ at a point a and continuity of $f(x)$ in a neighborhood of a suffice for the existence of $f'(a)$.

6.4 Generalizations of MVTs

The following result was proved by Reich (1969) which is the symmetric derivative version of a result established by Trahan (1966). This results generalizes the quasi-mean value theorem for functions with symmetric derivatives. For f differentiable on $[a, b]$, we adopt the convention $f'(b) = f^s(b)$.

Theorem 6.7 *Let f be continuous on the interval $[a, b]$ and symmetrically differentiable on the open interval $]a, b[$. Suppose f is differentiable at the right end point b of the interval $[a, b]$ and $[f(b) - f(a)] f'(b) \leq 0$. Then there exist points η, ξ in $]a, b[$ such that*

$$f^s(\eta) \geq 0 \quad \text{and} \quad f^s(\xi) \leq 0.$$

Proof: If $f'(b) = 0$ then letting $\eta = b$ and $\xi = b$, we have $f^s(\eta) = 0$ and $f^s(\xi) = 0$ using the convention $f^s(b) = f'(b)$.

If $f(b) = f(a)$, then applying Theorem 6.3 to f on $[a, b]$, we obtain

$$f^s(\xi) \leq 0 \leq f^s(\eta)$$

for some η and ξ in $]a, b[$.

Suppose $[f(b) - f(a)] f'(b) < 0$. This implies that either $f'(b) < 0$ and $f(b) > f(a)$ or $f'(b) > 0$ and $f(b) < f(a)$. In the first case, since f is continuous on $[a, b]$ and $f(b) > f(a)$ with f decreasing at b , there exists a point y in $]a, b[$ such that

$$f(y) > f(b) > f(a).$$

Hence applying Lemma 6.1 to f on the interval $[y, b]$, we obtain $f^s(\xi) \leq 0$ for some $\xi \in]y, b[$. Again applying Lemma 6.1 to f on the interval $[a, b]$, we get $f^s(\eta) \geq 0$ for some $\eta \in]a, b[$. The case when $f'(b) > 0$ and $f(b) < f(a)$ can be handled in a similar manner. This completes the proof of the theorem.

In the next theorem, we present a version of Flett's mean value theorem for symmetrically differentiable functions.

Theorem 6.8 *Let f be continuous on the interval $[a, b]$ and symmetrically differentiable on the open interval $]a, b[$. Suppose f is differentiable at the end points a and b of the interval $[a, b]$ and*

$$\left[f'(b) - \frac{f(b) - f(a)}{b - a} \right] \left[f'(a) - \frac{f(b) - f(a)}{b - a} \right] \geq 0.$$

Then there are points η, ξ in $]a, b]$ such that

$$(\eta - a) f^s(\eta) \geq f(\eta) - f(a)$$

and

$$(\xi - a) f^s(\xi) \leq f(\xi) - f(a).$$

Proof: Define $h : [a, b] \rightarrow \mathbb{R}$ by

$$h(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{if } x \in]a, b] \\ f'(a) & \text{if } x = a. \end{cases}$$

Then evidently, h is continuous on $[a, b]$ and symmetrically differentiable on $]a, b]$. Further, we have

$$h^s(x) = -\frac{h(x)}{x-a} + \frac{f^s(x)}{x-a}$$

for all $x \in]a, b]$. In view of

$$\left[f'(b) - \frac{f(b) - f(a)}{b-a} \right] \left[f'(a) - \frac{f(b) - f(a)}{b-a} \right] \geq 0,$$

we see that $[h(b) - h(a)] h'(b) \leq 0$. By above theorem, we obtain

$$h^s(\xi) \leq 0 \leq h^s(\eta)$$

for some $\xi, \eta \in]a, b]$. By the definition of h this amounts to the following inequalities:

$$(\eta - a) f^s(\eta) \geq f(\eta) - f(a)$$

and

$$(\xi - a) f^s(\xi) \leq f(\xi) - f(a).$$

This completes the proof of the theorem.

The above theorem can be further improved to have a Cauchy type mean value theorem for functions with symmetric derivatives.

Theorem 6.9 *Let f and g be continuous on the interval $[a, b]$ and symmetrically differentiable on the open interval $]a, b[$. Further, let f and g*

be both differentiable at the end points a and b of the interval $[a, b]$ with $g'(a) \neq 0 \neq g'(b)$. Suppose $g(x) \neq g(a)$ for all $x \in]a, b]$ and

$$\left[\frac{f'(b)}{g'(b)} - \frac{f(b) - f(a)}{g(b) - g(a)} \right] \left[\frac{f'(a)}{g'(a)} - \frac{f(b) - f(a)}{g(b) - g(a)} \right] \geq 0.$$

Then exist points η, ξ in $]a, b]$ such that

$$[g(\eta) - g(a)] f^s(\eta) \geq [f(\eta) - f(a)] g^s(\eta)$$

and

$$[g(\xi) - g(a)] f^s(\xi) \leq [f(\xi) - f(a)] g^s(\xi).$$

Proof: Define $h : [a, b] \rightarrow \mathbb{R}$ by

$$h(x) = \begin{cases} \frac{f(x) - f(a)}{g(x) - g(a)} & \text{if } x \in]a, b] \\ \frac{f'(a)}{g'(a)} & \text{if } x = a \end{cases}$$

and proceed in a manner similar to the proof of Theorem 6.8.

6.5 Dini Derivatives of Real Functions

The importance of Dini derivatives was felt in the last few decades when nonsmooth optimization problems arose in economics and engineering. The typical feature of Dini derivatives is that they always exist and admit very useful calculus rules, as well. This section is concerned with the Dini derivatives of a real valued function of one variable. We present only some of their basic properties, a comprehensive account of these derivatives can be found in the excellent paper by Giorgi and Komlosi (1992). We begin this section with the definition of Dini derivatives and then provide some examples for illustration.

Definition 6.3 Let $f :]a, b[\rightarrow \mathbb{R}$ be a real valued function and let $c \in]a, b[$ be an arbitrary point. The four *Dini derivatives* of f at c are defined as follows:

$$f_+(c) = \sup_{\delta > 0} \inf_{0 < x - c < \delta} \frac{f(x) - f(c)}{x - c}, \quad f_-(c) = \sup_{\delta > 0} \inf_{-\delta < x - c < 0} \frac{f(x) - f(c)}{x - c},$$

$$f^+(c) = \inf_{\delta > 0} \sup_{0 < x - c < \delta} \frac{f(x) - f(c)}{x - c}, \quad f^-(c) = \inf_{\delta > 0} \sup_{-\delta < x - c < 0} \frac{f(x) - f(c)}{x - c},$$

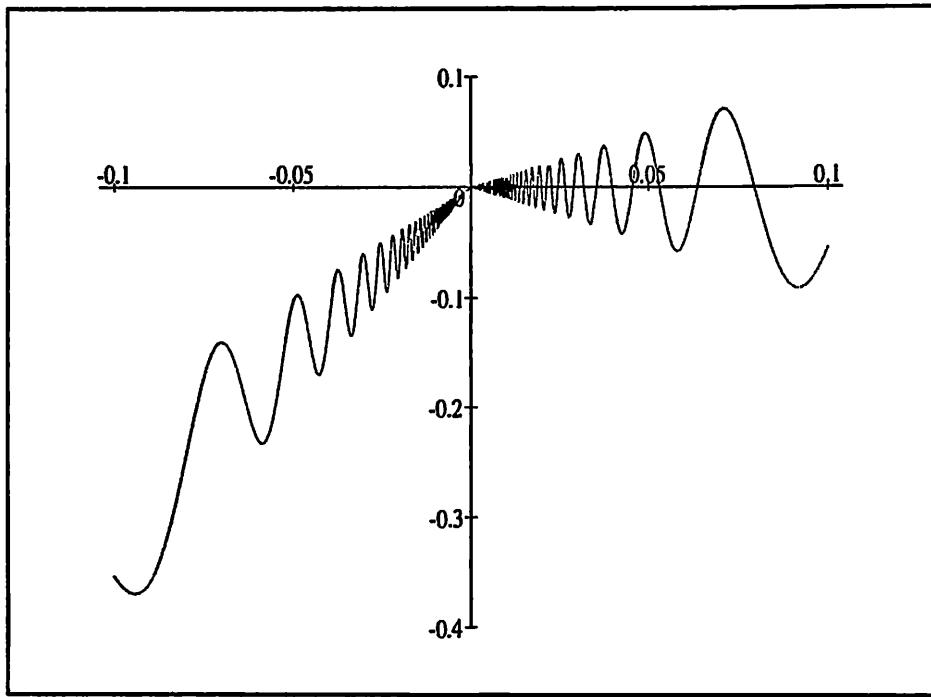


Fig. 6.4 The Graph of the Function $f(x)$ in Example 6.6.

where δ is a positive real number such that the neighborhood $|x - c| < \delta$ is contained in $]a, b[$. The quantities $f^+(c)$ and $f_+(c)$ are called the upper and lower right Dini derivative of f at c while $f^-(c)$ and $f_-(c)$ are called the upper and lower left Dini derivative of f at c .

Example 6.6 Consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} x \sin\left(\frac{1}{x}\right) & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ x [3 + \sin\left(\frac{1}{x}\right)] & \text{if } x < 0. \end{cases}$$

The Dini derivatives of f are given by $f_+(0) = -1$, $f_-(0) = 2$, $f^+(0) = 1$ and $f^-(0) = 4$.

From the definition of Dini derivatives, we get

$$f_+(0) = \sup_{\delta > 0} \inf_{0 < x < \delta} \frac{f(x) - f(0)}{x - 0} = \sup_{\delta > 0} \inf_{0 < x < \delta} \sin\left(\frac{1}{x}\right) = -1,$$

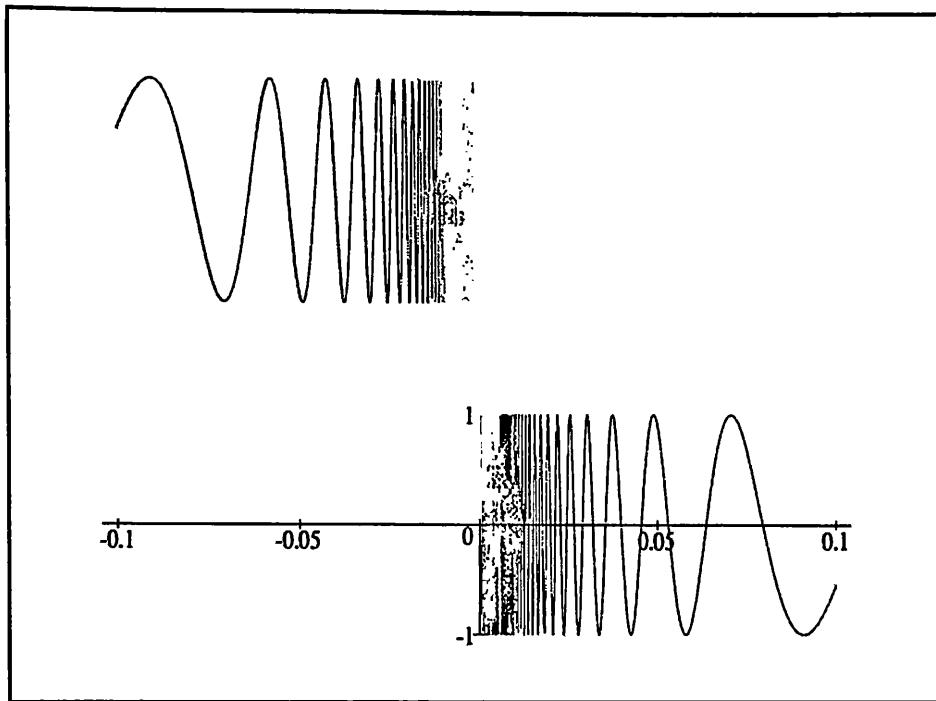


Fig. 6.5 The Graph of the Function $\frac{f(x)-f(0)}{x}$.

$$f_-(0) = \sup_{\delta > 0} \inf_{-\delta < x < 0} \frac{f(x) - f(0)}{x - 0} = \sup_{\delta > 0} \inf_{-\delta < x < 0} \left[3 + \sin\left(\frac{1}{x}\right) \right] = 2,$$

$$f^+(0) = \inf_{\delta > 0} \sup_{0 < x < \delta} \frac{f(x) - f(0)}{x - 0} = \sup_{\delta > 0} \inf_{0 < x < \delta} \sin\left(\frac{1}{x}\right) = 1,$$

and

$$f^-(0) = \inf_{\delta > 0} \sup_{-\delta < x < 0} \frac{f(x) - f(0)}{x - 0} = \sup_{\delta > 0} \inf_{-\delta < x < 0} \left[3 + \sin\left(\frac{1}{x}\right) \right] = 4.$$

The graph of $f(x)$ and $\frac{f(x)-f(0)}{x}$ are shown in Figure 6.4 and Figure 6.5, respectively. From these graphs, it is easy to visualize the Dini derivatives of f . Now we describe some rules for Dini derivatives by using the corresponding rules for *limsup* and *liminf*.

From the definition of Dini derivatives, it is easy to note that

$$f_+(c) \leq f^+(c) \quad \text{and} \quad f_-(c) \leq f^-(c).$$

A function f is differentiable at c if and only if all the four Dini derivatives are equal, that is

$$f'(c) = f_+(c) = f^+(c) = f_-(c) = f^-(c).$$

Non differentiable functions with finite Dini derivatives are closely related to differentiable functions: they are continuous and almost everywhere differentiable.

Theorem 6.10 *Let $f : [a, b] \rightarrow \mathbb{R}$ and c be a point in $]a, b[$. If each of the four Dini derivatives at c is finite, then the difference quotient*

$$\frac{f(x) - f(c)}{x - c}$$

is bounded around c , and hence f is continuous at c .

Proof: Suppose f is not continuous at c . Then there exists an $\epsilon > 0$ and a sequence $\{c_n\}$ converging to c such that

$$|f(c_n) - f(c)| > \epsilon.$$

Hence

$$\left| \frac{f(c_n) - f(c)}{c_n - c} \right| \rightarrow \infty$$

as $c_n \rightarrow c$. This implies that at least one of the four Dini derivatives is not finite which is a contradiction.

We state the following theorem without giving a proof. The interested reader should see Young (1916) for a proof.

Theorem 6.11 *Let f be a real valued function defined on the interval $[a, b]$. If all four Dini derivatives are finite at each point of the interval $[a, b]$, then f is differentiable almost everywhere.*

The following example shows that if all four Dini derivatives are not finite, continuity is in general not ensured.

Example 6.7 Consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2 & \text{if } x \geq 0 \\ 1 + x & \text{if } x < 0. \end{cases}$$

Then Dini derivatives of f are given by $f_+(0) = 0$, $f_-(0) = -\infty$, $f^+(0) = 0$ and $f^-(0) = \infty$.

This can be seen from the definition of Dini derivatives. For example,

$$\begin{aligned} f^+(0) &= \inf_{\delta>0} \sup\{x \mid 0 < x < \delta\} \\ &= \inf_{\delta>0} \delta \\ &= 0, \end{aligned}$$

$$\begin{aligned} f_+(0) &= \sup_{\delta>0} \inf\{x \mid 0 < x < \delta\} \\ &= \inf_{\delta>0} \delta \\ &= 0, \end{aligned}$$

$$\begin{aligned} f^-(0) &= \inf_{\delta>0} \sup\left\{1 + \frac{1}{x} \mid -\delta < x < 0\right\} \\ &= \inf_{\delta>0} \infty \\ &= \infty, \end{aligned}$$

$$\begin{aligned} f^-(0) &= \sup_{\delta>0} \inf\left\{1 + \frac{1}{x} \mid -\delta < x < 0\right\} \\ &= \inf_{\delta>0} -\infty \\ &= -\infty. \end{aligned}$$

Definition 6.4 A real valued function f is said to be *locally strictly increasing* at c if there exists an open interval $]c - \delta, c + \delta[$ such that

$$f(x) < f(c) \quad \text{for all } c - \delta < x < c$$

and

$$f(x) > f(c) \quad \text{for all } c < x < c + \delta.$$

The following two results will be instrumental in establishing Rolle's Theorem and the mean value theorem for nondifferentiable functions.

Theorem 6.12 Let $f : [a, b] \rightarrow \mathbb{R}$ be real valued function defined on the interval $[a, b]$ and c be a point in $]a, b[$. If $f_+(c) > 0$ and $f_-(c) > 0$, then f is locally strictly increasing at c .

Proof: Suppose $f_+(c) > 0$ and $f_-(c) > 0$. Then there exists $x_1, x_2 \in [a, b]$ with $x_1 < c < x_2$ such that

$$\frac{f(x) - f(c)}{x - c} > 0$$

for every $x \in]x_1, x_2[$, $x \neq c$. Hence, we have

$$f(x) < f(c) \quad \text{if } x_1 < x < c$$

and

$$f(x) > f(c) \quad \text{if } c < x < x_2.$$

This proves that f is locally strictly increasing at c . This completes the proof of the theorem.

Similarly, if $f^+(c) < 0$ and $f^-(c) < 0$, then f is locally strictly decreasing at the point c . The proof of this statement is analogous to the proof of the above theorem.

Theorem 6.13 *Let $f : [a, b] \rightarrow \mathbb{R}$ be real valued function defined on the interval $[a, b]$. If f is locally strictly increasing at every point of $[a, b]$, then f is strictly increasing on $[a, b]$.*

Proof: Let f be locally strictly increasing at each point of the interval $[a, b]$. We want to prove that f is strictly increasing on $[a, b]$. Suppose, f is not strictly increasing on $[a, b]$. Then there exist $x_1, x_2 \in [a, b]$, such $x_1 < x_2$ and

$$f(x_1) \geq f(x_2).$$

We define a set T by

$$T = \{x \in]x_1, x_2[\mid f(x_1) \leq f(x)\}.$$

Since f is locally strictly increasing at x_1 and x_2 , therefore $T \neq \emptyset$. Let

$$x_o = \sup T.$$

Then clearly $x_1 < x_o < x_2$, since f is strictly locally increasing at x_1 and x_2 . Again using the fact that f is locally strictly increasing at x_o , we get

$$f(x) < f(x_o) \quad \text{for } x_o - \delta < x < x_o \tag{6.6}$$

and

$$f(x) > f(x_o) \quad \text{for } x_o < x < x_o + \delta \quad (6.7)$$

for some $\delta > 0$. By the definition of x_o , if $x_o < x < x_2$ then $x \notin T$ and hence

$$f(x) < f(x_1). \quad (6.8)$$

Using (6.7) and (6.8), we see that $f(x_o) < f(x) < f(x_1)$ for any $x_o < x < x_o + \delta$, thus

$$f(x_o) < f(x_1).$$

On the other hand there exists x' in $x_o - \delta < x' < x_o$ such that $x' \in T$. Hence by (6.6) and the definition of T , one has

$$f(x_1) \leq f(x') < f(x_o).$$

This contradicts the fact that $f(x_1) > f(x_o)$. This contradiction proves that f is strictly increasing on $[a, b]$ and the proof of the theorem is now complete.

6.6 MVTs for Nondifferentiable Functions

In this section, we study the mean value theorem and its various generalizations for nondifferentiable functions using Dini derivatives. Our first theorem is due to Castagnoli (1983). In this theorem no extra conditions are imposed on the function.

Theorem 6.14 *Let $f : [a, b] \rightarrow \mathbb{R}$ be real valued function with $f(a) = f(b)$. Then there exist $\eta, \xi \in [a, b]$ such that*

$$\text{either } f_+(\eta) \leq 0 \quad \text{or} \quad f_-(\eta) \leq 0 \quad (6.9)$$

and

$$\text{either } f^+(\xi) \geq 0 \quad \text{or} \quad f^-(\xi) \geq 0. \quad (6.10)$$

Proof: We shall establish (6.9) and leave the proof of (6.10) to the reader as the proof is similar to the proof of (6.9).

Suppose there does not exist an $\eta \in [a, b]$ such that either $f_+(\eta) \leq 0$ or $f_-(\eta) \leq 0$. This implies that

$$f_+(x) > 0 \quad \text{and} \quad f_-(x) > 0$$

for all $x \in [a, b]$. By Theorem 6.12, f is locally strictly increasing at every point x in $[a, b]$. By Theorem 6.13, f is strictly increasing on $[a, b]$. But this is a contradiction to the fact that $f(a) = f(b)$. The proof is now complete.

If we impose some special conditions on the function we can obtain more information about the behavior of the Dini derivatives.

Theorem 6.15 *Let $f : [a, b] \rightarrow \mathbb{R}$ be a real valued function with $f(a) = f(b)$. Suppose f assumes a minimum m and a maximum M on $[a, b]$. If $m < f(a) = f(b)$, then there exists $\eta \in]a, b[$ such that*

$$f^-(\eta) \leq 0 \leq f_+(\eta).$$

If $M > f(a) = f(b)$, then there exists $\xi \in]a, b[$ such that

$$f^+(\xi) \leq 0 \leq f_-(\xi).$$

Proof: Let $\eta \in [a, b]$ such that $f(\eta) = m$. Suppose $m < f(a) = f(b)$, then $\eta \in]a, b[$. Since in this case we have

$$f(x) \geq f(\eta)$$

for all $x \in [a, b]$, therefore

$$f^-(\eta) \leq 0 \leq f_+(\eta)$$

holds. The proof of the other part of the theorem follows by similar arguments and the proof of the theorem is now complete.

Theorem 6.16 *Let $f : [a, b] \rightarrow \mathbb{R}$ be a real valued continuous function with $f(a) = f(b)$. Then there exist intermediate values $\eta_1, \eta_2, \eta_3, \eta_4 \in]a, b[$ such that*

$$f_+(\eta_1) \geq 0, \quad f^+(\eta_2) \leq 0, \quad f_-(\eta_3) \geq 0, \quad f^-(\eta_4) \leq 0.$$

Proof: Since f is continuous, the image of $[a, b]$ is a closed interval:

$$\{f(x) \mid x \in [a, b]\} = [m, M].$$

Suppose $m = M$. In this case f is constant over $[a, b]$ so all of the Dini derivatives are equal to 0 for every $x \in]a, b[$ and the theorem holds.

Suppose $M > f(a)$. Then, there exists $d \in]a, b[$ such that $f(d) = M$. Let $f(a) < y < M$. Since f is continuous there are two points $d_1 \in]a, d[$ and $d_2 \in]d, b[$ such that f assumes the value y . That is

$$f(d_1) = y = f(d_2).$$

Put $\eta_1 = d_1$, $\eta_2 = d$, $\eta_3 = d$, $\eta_4 = d_2$. It is not difficult to show that the inequalities are fulfilled with this choice.

Suppose $m < f(a)$. The proof goes by the same arguments of the previous case. The proof of the theorem is now complete.

The following theorem can be easily deduced from the above theorem and we leave its proof to the reader.

Theorem 6.17 *Let $f : [a, b] \rightarrow \mathbb{R}$ be a real valued continuous function. Then there exist intermediate values $\eta_1, \eta_2, \eta_3, \eta_4 \in]a, b[$ such that*

$$f_+(\eta_1) \geq \gamma, \quad f^+(\eta_2) \leq \gamma, \quad f_-(\eta_3) \geq \gamma, \quad f^-(\eta_4) \leq \gamma,$$

where $\gamma = \frac{f(b) - f(a)}{b - a}$.

Next, we present an extension of the Flett type mean value theorem due to Lakshminarasimhan (1966).

Theorem 6.18 *Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous on $[a, b]$ and differentiable at $x = a$ and $x = b$. Further, let the four Dini derivatives of f be finite in $]a, b[$. Then if $f'(a) = f'(b)$ there exists a point $\eta \in]a, b[$, such that*

$$f^+(\eta) \leq \frac{f(\eta) - f(a)}{\eta - a} \leq f_-(\eta)$$

or a point $\xi \in]a, b[$, such that

$$f^-(\xi) \leq \frac{f(\xi) - f(a)}{\xi - a} \leq f_+(\xi).$$

Proof: Without loss of generality, let us assume that $f'(a) = f'(b) = 0$. For otherwise we have to consider only $f(x) - xf'(a)$. Define $g : [a, b] \rightarrow \mathbb{R}$ by

$$g(x) = \begin{cases} \frac{f(x) - f(a)}{x - a} & \text{if } x \in]a, b[\\ f'(a) & \text{if } x = a. \end{cases}$$

Evidently, the function g is continuous in $[a, b]$. Further

$$g'(b) = -\frac{g(b)}{b-a}.$$

From the above equality, we see that if $g(b) > 0$, then $g'(b) < 0$. Hence g is a decreasing function at b , while $g(a) = 0$. Since g is continuous in $[a, b]$, it attains maximum at a point $\eta \in]a, b[$. Hence

$$g^+(\eta) \leq 0 \quad \text{and} \quad g_-(\eta) \geq 0.$$

But

$$g^+(\eta) = -\frac{f(\eta) - f(a)}{(\eta - a)^2} + \frac{f^+(\eta)}{\eta - a}$$

and

$$g_-(\eta) = -\frac{f(\eta) - f(a)}{(\eta - a)^2} + \frac{f_-(\eta)}{\eta - a}.$$

Hence we have the inequalities

$$f^+(\eta) \leq \frac{f(\eta) - f(a)}{\eta - a} \leq f_-(\eta).$$

If, on the other hand, $g(b) < 0$, then $g'(b) > 0$, so that g is increasing at b while $g(a) = 0$. Hence, g attains its minimum at a point $\xi \in]a, b[$; so that

$$g_+(\xi) \geq 0 \quad \text{and} \quad g^-(\xi) \leq 0,$$

and we obtain the other inequalities.

Finally, if $g(b) = 0$, then g is continuous in the closed interval $[a, b]$ and $g(a) = 0$, g attains either a maximum at a point η or a minimum at a point ξ between a and b . Hence, as before, either $g^+(\eta) \leq 0$ and $g_-(\eta) \geq 0$ or $g_+(\xi) \geq 0$ and $g^-(\xi) \leq 0$. These inequalities yield the asserted inequalities of the theorem and the proof is now complete.

The following result was established by Reich (1969) which is analogous of a result proved by Trahan (1966).

Theorem 6.19 *Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous in $[a, b]$ and differentiable at the right end point b . Further, let the four Dini derivatives of f be finite in $]a, b[$ and let $[f(b) - f(a)] f'(b) \leq 0$. Then there exist points η, ξ in $]a, b[$ such that*

$$f^+(\eta) \leq 0 \leq f_-(\eta)$$

or

$$f^-(\xi) \leq 0 \leq f_+(\xi).$$

Proof: If $f'(b) = 0$, then letting $\eta = b$ and $\xi = b$, we have the asserted inequalities of the theorem.

If $f(b) = f(a)$, and f is not constant then by continuity, f attains its maximum at η and its minimum at ξ in $]a, b[$. Thus we obtain the asserted inequalities of the theorem.

Next, suppose $[f(b) - f(a)]f'(b) < 0$. This implies that either $f'(b) < 0$ and $f(b) > f(a)$ or $f'(b) > 0$ and $f(b) < f(a)$. In the first case, since f is continuous on $[a, b]$ and $f(b) > f(a)$ with f decreasing at b , the function f has a maximum at $\eta \in]a, b[$. Hence we obtain the first set of inequalities of the theorem. Similarly, in the second case f has a minimum at some point $\xi \in]a, b[$ and hence the second set of inequalities of the theorem follows. This completes the proof of the theorem.

Theorem 6.20 *Let f be continuous on the interval $[a, b]$ and differentiable at the points a and b . Further, suppose the four Dini derivatives are finite in $]a, b[$ and*

$$\left[f'(b) - \frac{f(b) - f(a)}{b - a} \right] \left[f'(a) - \frac{f(b) - f(a)}{b - a} \right] \geq 0.$$

Then there exist points η, ξ in $]a, b]$ such that

$$f^+(\eta) \leq \frac{f(\eta) - f(a)}{\eta - a} \leq f_-(\eta)$$

or

$$f^-(\xi) \leq \frac{f(\xi) - f(a)}{\xi - a} \leq f_+(\xi).$$

Proof: Define $h : [a, b] \rightarrow \mathbb{R}$ by

$$h(x) = \begin{cases} \frac{f(x) - f(a)}{x - a} & \text{if } x \in]a, b] \\ f'(a) & \text{if } x = a. \end{cases}$$

Then evidently, h satisfies the conditions of the previous theorem and hence this theorem follows.

In addition to the above, a Cauchy type mean value theorem and its generalizations have been studied. Since the development of these theorems are quite parallel to the usual Cauchy type mean value theorem, we shall

not pursue these developments here and rather leave them to reader as an exercise.

In this chapter, we examined Rolle's theorem and Lagrange's mean value theorem and their generalizations for functions with symmetric derivatives and Dini derivatives. There are many other generalized derivatives such as approximate derivative, preponderant derivative, qualitative derivative, congruent derivative, selective derivative, path derivative, Csaszar's derivative, and Garg's derivative (see Bruckner (1994)). For some of these generalized derivatives some results related to Lagrange's mean value theorem exist and other results can be established. Because of the introductory nature of this book we refrain from mean value type results for these generalized derivatives.

Chapter 7

Some Integral Mean Value Theorems and Related Topics

The mean value theorem of differential calculus is very important as we have seen in Chapter two. Another important theorem in calculus is the integral mean value theorem. The main goal of this chapter is to present the integral mean value theorem and some generalizations of this theorem. Among others, we present generalizations such as Bonnet's mean value theorem and Sayrafiezadeh's mean value theorem. We also present a result due to Wayment (1970) which is the integral version of Flett's mean value theorem. Using the integral mean value theorem we present integral representations of several known means such as arithmetic, geometric, logarithmic and identric. Further, in this chapter, we investigate the nature of the function f which in some sense connects the mean value theorem of differential calculus to that of integral calculus. Finally, some open problems related to the iteration of means that need an investigation will be discussed.

7.1 The Integral MVT and Generalizations

The mean value theorem for integrals is established using the second fundamental theorem of calculus, which states that if $f(x)$ is continuous on an interval $[a, b]$ and $F(x) = \int_a^x f(t) dt$ for $x \in [a, b]$, then $F'(x) = f(x)$ for all $x \in]a, b[$. This result is attributed to Isaac Barrow (1630-1677) who was the first to realize that differentiation and integration are inverse operations. The mean value theorem of integral calculus states that

Theorem 7.1 *If $f(x)$ is continuous on $[x, y]$, then there exists a point ξ*

in $]x, y[$ depending on x and y such that

$$f(\xi(x, y)) = \frac{\int_x^y f(t) dt}{y - x}. \quad (7.1)$$

Proof: Let us define

$$F(z) = \int_a^z f(t) dt,$$

where a is a constant in $]x, y[$ and $z \in]x, y[$. Since f is continuous on the interval $[x, y]$, the function F is also continuous on $[x, y]$ and F is also differentiable on the open interval $]x, y[$. By the second fundamental theorem of calculus, we have $F'(x) = f(x)$. Now we apply the mean value theorem of differential calculus to the function F . Then there exists a point ξ in the interval $]x, y[$ depending on x and y such that

$$\frac{F(y) - F(x)}{y - x} = F'(\xi(x, y)),$$

that is

$$\frac{\int_a^y f(t) dt - \int_a^x f(t) dt}{y - x} = f(\xi(x, y)).$$

Hence, we have the asserted statement

$$f(\xi(x, y)) = \frac{\int_x^y f(t) dt}{y - x}$$

and the proof of the theorem is now complete.

A generalization of the mean value theorem of integral calculus is the following:

Theorem 7.2 *If f and g are continuous on $[a, b]$ and g is never zero on $]a, b[$, then there exists a number ξ in $]a, b[$ depending on a and b such that*

$$f(\xi(a, b)) = \frac{\int_a^b g(t) f(t) dt}{\int_a^b g(t) dt}. \quad (7.2)$$

Proof: We define two functions $H(x)$ and $G(x)$ in $[a, b]$ by

$$H(x) = \int_a^x f(t) g(t) dt$$

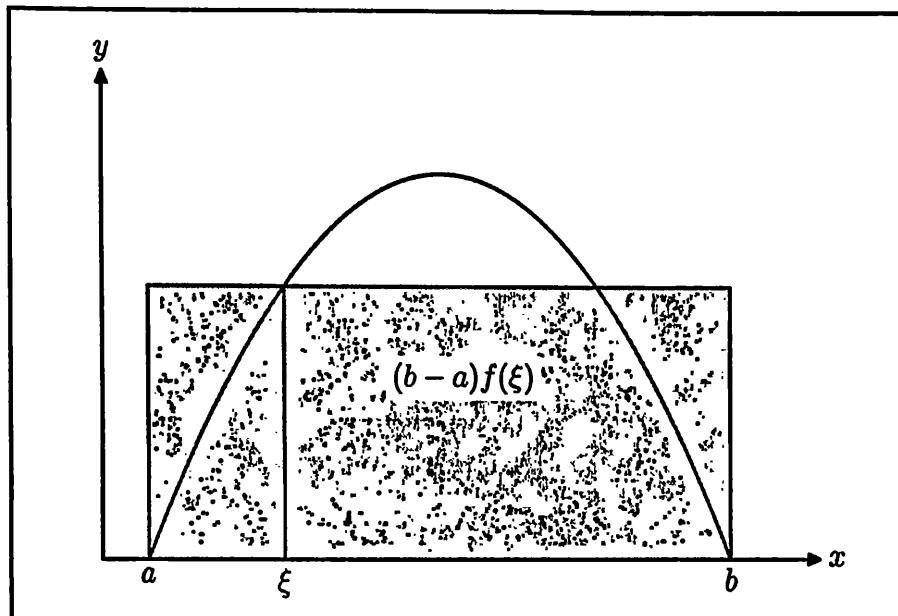


Fig. 7.1 A Geometrical Illustration of the Integral Mean Value Theorem.

$$G(x) = \int_a^x g(t) dt.$$

Consider the function

$$D(x) = \begin{vmatrix} H(x) & G(x) & 1 \\ H(a) & G(a) & 1 \\ H(b) & G(b) & 1 \end{vmatrix}.$$

Since $G(a) = 0 = H(a)$, we get

$$D(x) = \begin{vmatrix} H(x) & G(x) & 1 \\ 0 & 0 & 1 \\ H(b) & G(b) & 1 \end{vmatrix}.$$

Note that $D(a) = 0 = D(b)$. Applying Rolle's theorem to $D(x)$ on the interval $[a, b]$, we get

$$D'(\xi) = 0$$

for some $\xi \in]a, b[$. That is

$$\begin{aligned} 0 &= D'(\xi) \\ &= \begin{vmatrix} H'(\xi) & G'(\xi) & 1 \\ 0 & 0 & 1 \\ H(b) & G(b) & 1 \end{vmatrix} = \begin{vmatrix} f(\xi)g(\xi) & g(\xi) & 0 \\ 0 & 0 & 1 \\ H(b) & G(b) & 1 \end{vmatrix} \\ &= g(\xi) [-f(\xi)G(b) + H(b)]. \end{aligned}$$

Hence

$$f(\xi) = \frac{H(b)}{G(b)}$$

that is

$$f(\xi(a, b)) = \frac{\int_a^b g(t) f(t) dt}{\int_a^b g(t) dt}.$$

This completes the proof.

Another generalization of the integral mean value theorem was given by Wayment (1970). We present his result in the next theorem. The following integral mean value theorem can be considered analogous to Flett's mean value theorem of differential calculus.

Theorem 7.3 *If f is continuous on the closed interval $[a, b]$ and $f(a) = f(b)$, then there exists a number η in $]a, b[$ depending on a and b such that*

$$(\eta - a) f(\eta) = \int_a^\eta f(x) dx. \quad (7.3)$$

Proof: Define $F : [a, b] \rightarrow \mathbb{R}$ by

$$F(t) = (t - a) f(t) - \int_a^t f(x) dx \quad (7.4)$$

for all t in $[a, b]$. If f is a constant function, then $F = 0$ and the conclusion of the theorem holds. Hence, we assume from now on that f is not a constant function. Since f is continuous on the closed and bounded interval $[a, b]$, it has a maximum and a minimum on $[a, b]$. Hence there exist t_1 and t_2 in $[a, b]$ such that

$$f(t_1) \leq f(t) \leq f(t_2). \quad (7.5)$$

Since $f(a) = f(b)$, at least one of them is not the left end point a of the interval $[a, b]$. Suppose $t_2 \neq a$. Then $t_2 \neq b$, otherwise f becomes a constant function. In this situation, we have

$$F(b) < 0 < F(t_2).$$

Hence, by an application of the intermediate value theorem, we have

$$F(\eta) = 0,$$

that is

$$(\eta - a) f(\eta) = \int_a^\eta f(x) dx$$

for some $\eta \in]t_2, b[$.

Next, suppose $t_1 \neq a$. Then t_1 cannot be equal to b either. Further, we have

$$F(t_1) < 0 < F(b).$$

Hence, as before, we obtain

$$F(\eta) = 0,$$

that is

$$(\eta - a) f(\xi) = \int_a^\eta f(x) dx$$

for some $\eta \in]t_1, b[$.

If t_1 and t_2 are both not equal to a . Then they can not also be equal to b . Therefore $a < \min\{t_1, t_2\} < \max\{t_1, t_2\} < b$ and

$$F(t_1) \leq 0 \leq F(t_2).$$

Hence by the intermediate value theorem, we get

$$F(\eta) = 0,$$

that is

$$(\eta - a) f(\eta) = \int_a^\eta f(x) dx$$

for some $\eta \in [t_1, t_2]$. From the above cases we conclude that (7.3) holds for $\eta \in]a, b[$. The proof is now complete.

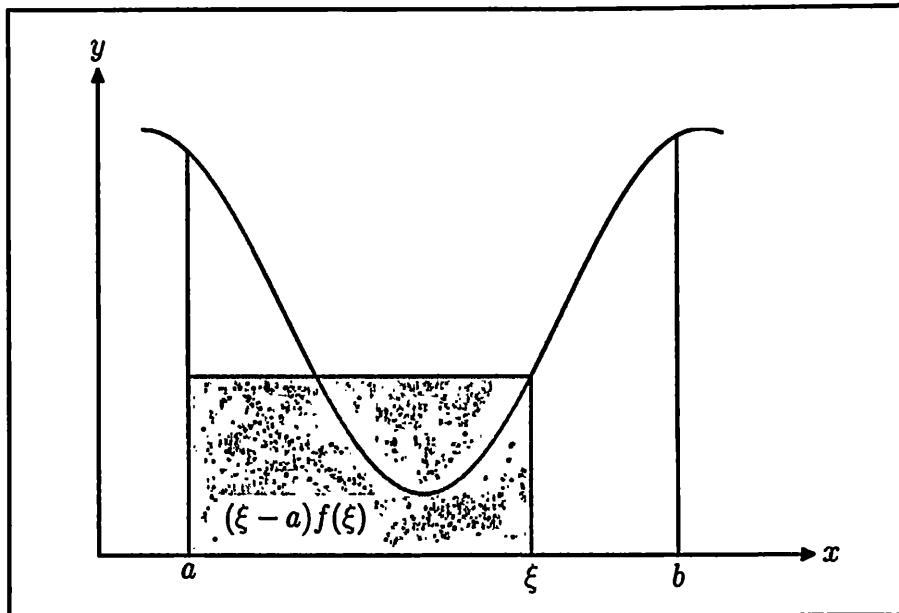


Fig. 7.2 An Illustration of the Integral Mean Value Theorem of Flett Type.

Now we present another generalization of the integral mean value theorem. This generalization is due to Sayrafiezadeh (1995).

Definition 7.1 Let $f : [a, b] \rightarrow \mathbb{R}$ be a continuous function. If $a = x_0 < x_1 < x_2 < \dots < x_n = b$, then the set

$$\mathcal{P}_n = \{x_0, x_1, \dots, x_n\}$$

is called a *partition* of the interval $[a, b]$.

Definition 7.2 Given a partition $\mathcal{P}_n = \{x_0, x_1, \dots, x_n\}$ of $[a, b]$, the norm of \mathcal{P}_n , denoted by $\|\mathcal{P}_n\|$, is defined as the length of the subinterval of maximum length.

Definition 7.3 Given a partition $\mathcal{P}_n = \{x_0, x_1, \dots, x_n\}$ of $[a, b]$, let $c_i \in [x_{i-1}, x_i]$ for $i = 1, 2, \dots, n$. The Riemann sum for f over $[a, b]$ is defined as

$$\mathcal{R}_n = \sum_{i=1}^n f(c_i) \Delta x_i$$

where $\Delta x_i = x_i - x_{i-1}$.

In the above Riemann sum, there are many ways of choosing the “sample points” c_1, c_2, \dots, c_n . We will choose these sample points in such a way that

they divide the corresponding subintervals in a fixed proportion. In other words, for a partition $\mathcal{P}_n = \{x_0, x_1, \dots, x_n\}$ there is a fixed number t in $[0, 1]$ such that

$$c_i = x_{i-1} + t \Delta x_i$$

for $i = 1, 2, \dots, n$. If we choose the sample points c_1, c_2, \dots, c_n in this manner, then the above Riemann sum becomes a function of t for a fixed partition and we denote this by

$$\begin{aligned}\mathcal{R}_n(t) &= \sum_{i=1}^n f(c_i) \Delta x_i \\ &= \sum_{i=1}^n f(x_{i-1} + t \Delta x_i) \Delta x_i.\end{aligned}$$

Notice that by choosing the different values for t , we obtain the various Riemann sums. For example, if $t = 0$, the $\mathcal{R}_n(0)$ represents the left endpoint Riemann sum. Similarly $\mathcal{R}_n(\frac{1}{2})$ and $\mathcal{R}_n(1)$ represent the midpoint and right endpoint Riemann sums, respectively.

In the following theorem, we address the question: For a given partition does there exist a point $\eta \in]0, 1[$ such that $\mathcal{R}_n(\eta) = \int_a^b f(x) dx$. In the case $n = 1$, the following theorem yields the integral mean value theorem. Thus in this sense this theorem is a generalization of the integral mean value theorem.

Theorem 7.4 *Suppose $f : [a, b] \rightarrow \mathbb{R}$ is a continuous function. Let $\mathcal{P}_n = \{x_0, x_1, \dots, x_n\}$ be a fixed partition of $[a, b]$. For $t \in [0, 1]$, let the sample points be given by $c_i = x_{i-1} + t \Delta x_i$ where $\Delta x_i = x_i - x_{i-1}$. Then there exists a point $\eta \in]0, 1[$ such that*

$$\mathcal{R}_n(\eta) = \int_a^b f(x) dx.$$

Proof: If $n = 1$, then this is just the integral mean value theorem. Suppose $n \geq 1$. The function $\mathcal{R}_n(t)$ is continuous on the interval $[0, 1]$ since the functions $f(x_{i-1} + t \Delta x_i)$ are composites of the continuous function $f(x)$ with the continuous function $x_{i-1} + t \Delta x_i$. Hence, the function $\mathcal{R}_n(t)$ is

integrable, and

$$\begin{aligned}
 \int_0^1 \mathcal{R}_n(t) dt &= \int_0^1 \left[\sum_{i=1}^n f(x_{i-1} + t \Delta x_i) \Delta x_i \right] dt \\
 &= \sum_{i=1}^n \int_0^1 f(x_{i-1} + t \Delta x_i) \Delta x_i dt \\
 &= \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(c_i) dc_i \\
 &= \int_a^b f(x) dx,
 \end{aligned}$$

where $c_i = x_{i-1} + t \Delta x_i$. Now applying the integral mean value theorem to the continuous function $\mathcal{R}_n(t)$ on the interval $[0, 1]$, we get

$$\mathcal{R}_n(\eta) = \int_0^1 \mathcal{R}_n(t) dt$$

for some $\eta \in]0, 1[$. Since the right hand integral is equal to $\int_a^b f(x) dx$, we obtain

$$\mathcal{R}_n(\eta) = \int_a^b f(x) dx$$

and the proof of the theorem is now complete.

We illustrate this theorem by an example. Let $f(x) = x^2$ for $0 \leq x \leq 3$. Let $\mathcal{P}_3 = \{0, 1, 2, 3\}$. Then $c_1 = t$, $c_2 = 1 + t$ and $c_3 = 2 + t$. Therefore

$$\mathcal{R}_3(t) = \sum_{i=1}^3 f(c_i) \Delta x_i = t^2 + (1+t)^2 + (2+t)^2.$$

Simplifying this we get $\mathcal{R}_3(t) = 3t^2 + 6t + 5$. Integrating $f(x) = x^2$ from 0 to 3, we obtain $\int_0^3 f(x) dx = 9$. Therefore, we have the quadratic equation $3t^2 + 6t + 5 = 9$. Solving this equation, we see that $\eta = 0.5276$. The Figure 7.3 illustrates this theorem geometrically with $n = 3$.

Now we present yet another generalization of the integral mean value theorem known as Bonnet's mean value theorem. As the name suggests, this theorem was discovered by Ossian Bonnet in 1849. Bonnet pointed out that his mean value theorem has many applications. However, we shall not discuss its applications here and refer the interested reader to the paper by Bonnet (1849) and the book by Bressoud (1994).

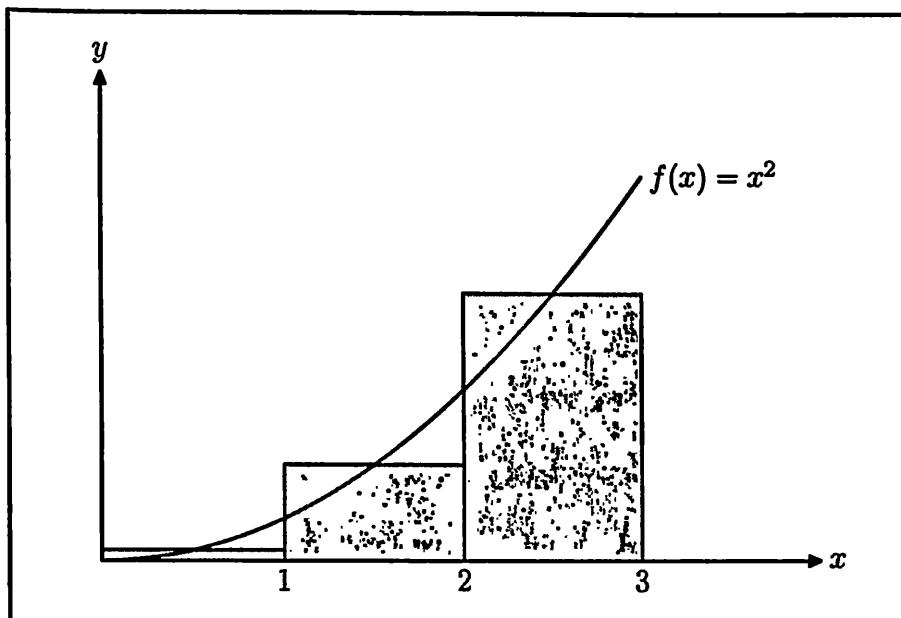


Fig. 7.3 An Illustration of Theorem 7.4.

Theorem 7.5 Let f be integrable and g be a nonnegative, increasing function on the closed interval $[a, b]$. Then there exists at least one point η in $]a, b[$ such that

$$\int_a^b f(t)g(t) dt = g(b) \int_\eta^b f(t) dt.$$

Proof: Let $\mathcal{P} = \{a = t_0, t_1, \dots, t_n = b\}$ be an arbitrary partition of the interval $[a, b]$. The integral $\int_a^b f(t)g(t) dt$ can be approximated by the right sum

$$S(\mathcal{P}) = \sum_{k=1}^n f(t_k)g(t_k)(t_k - t_{k-1}). \quad (7.6)$$

Let

$$S_m(\mathcal{P}) = \sum_{k=m+1}^n f(t_k)(t_k - t_{k-1}). \quad (7.7)$$

The partial sum S_m approximates the integral $\int_{x_m}^b f(t)dt$. Let A and B be the upper and lower bounds of the integral $\int_x^b f(t)dt$ for each x in $[a, b]$,

that is

$$A \leq \int_x^b f(t) dt \leq B \quad (7.8)$$

for each x in $[a, b]$. Let $A(\mathcal{P})$ and $B(\mathcal{P})$ be the lower and upper bounds of $\mathcal{S}_m(\mathcal{P})$ for all $m = 1, 2, \dots, n$. Hence

$$A(\mathcal{P}) \leq \mathcal{S}_m(\mathcal{P}) \leq B(\mathcal{P}) \quad (7.9)$$

for each m . From (7.7), we see that

$$f(t_k)(t_k - t_{k-1}) = \mathcal{S}_{k-1}(\mathcal{P}) - \mathcal{S}_k(\mathcal{P}). \quad (7.10)$$

Using (7.10) in (7.6) and simplifying, we get

$$\begin{aligned} \mathcal{S}(\mathcal{P}) &= g(t_1)\mathcal{S}_0(\mathcal{P}) + [g(t_2) - g(t_1)]\mathcal{S}_1(\mathcal{P}) + \cdots + \\ &\quad [g(t_{n-1}) - g(t_{n-2})]\mathcal{S}_{n-2}(\mathcal{P}) + g(b)\mathcal{S}_{n-1}(\mathcal{P}). \end{aligned} \quad (7.11)$$

Since, g is nonnegative and increasing on $[a, b]$, we obtain from (7.11)

$$\begin{aligned} \mathcal{S}(\mathcal{P}) &\leq g(t_1)B(\mathcal{P}) + [g(t_2) - g(t_1)]B(\mathcal{P}) + \cdots + \\ &\quad [g(t_{n-1}) - g(t_{n-2})]B(\mathcal{P}) + g(b)B(\mathcal{P}) \end{aligned} \quad (7.12)$$

which is

$$\mathcal{S}(\mathcal{P}) \leq g(b)B(\mathcal{P}).$$

Similarly, from (7.11), we also have

$$\begin{aligned} \mathcal{S}(\mathcal{P}) &\geq g(t_1)A(\mathcal{P}) + [g(t_2) - g(t_1)]A(\mathcal{P}) + \cdots + \\ &\quad [g(t_{n-1}) - g(t_{n-2})]A(\mathcal{P}) + g(b)A(\mathcal{P}) \end{aligned} \quad (7.13)$$

that is

$$\mathcal{S}(\mathcal{P}) \geq g(b)A(\mathcal{P}).$$

Therefore, we obtain

$$g(b)A(\mathcal{P}) \leq \mathcal{S}(\mathcal{P}) \leq g(b)B(\mathcal{P}).$$

As the partition becomes finer, the sum $S(\mathcal{P})$ tends to the integral $\int_a^b f(t)g(t) dt$, that is

$$\lim_{\|\mathcal{P}\| \rightarrow 0} S(\mathcal{P}) = \int_a^b f(t)g(t) dt$$

and

$$\lim_{\|\mathcal{P}\| \rightarrow 0} A(\mathcal{P}) = A \quad \text{and} \quad \lim_{\|\mathcal{P}\| \rightarrow 0} B(\mathcal{P}) = B.$$

Hence, we have

$$g(b) \lim_{\|\mathcal{P}\| \rightarrow 0} A(\mathcal{P}) \leq \lim_{\|\mathcal{P}\| \rightarrow 0} S(\mathcal{P}) \leq g(b) \lim_{\|\mathcal{P}\| \rightarrow 0} B(\mathcal{P}),$$

that is

$$Ag(b) \leq \int_a^b f(t)g(t) dt \leq Bg(b). \quad (7.14)$$

Now for $x \in [a, b]$, we define $\phi : [a, b] \rightarrow \mathbb{R}$ by

$$\phi(x) = g(b) \int_x^b f(t) dt.$$

Then, clearly ϕ is continuous on $[a, b]$ and $Ag(b) \leq \phi(x) \leq Bg(b)$. Applying the intermediate value theorem to ϕ , we have (in view of (7.14))

$$\phi(\eta) = \int_a^b f(t)g(t) dt$$

for some $\eta \in]a, b[$. This implies that there exists a point η in $]a, b[$ such that

$$g(b) \int_\eta^b f(t) dt = \int_a^b f(t)g(t) dt$$

and the proof of the theorem is now complete.

As an example, let $f(t) = \sin(t)$ and $g(x) = t^2$ on the interval $[0, 2\pi]$. Then by Bonnet's mean value theorem, we can find a number η between 0 and 2π for which

$$\int_0^{2\pi} t^2 \sin(t) dt = 4\pi^2 \int_\eta^{2\pi} \sin(t) dt = 4\pi^2 \cos(\eta).$$

7.2 Integral Representation of Means

We have seen that Lagrange's mean value theorem of differential calculus can be used for systematically generating various means between two positive real numbers a and b . In this section, we illustrate how the mean value theorem of integral calculus can be used for finding the integral representation of various means. These representations can be used to generalize the existing class of means. We begin this section with a formal definition of means.

Definition 7.4 A continuous function $M : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ is said to be a mean of two numbers a and b if and only if

- (M1) $\min\{a, b\} \leq M(a, b) \leq \max\{a, b\}$
- (M2) $M(a, b) = M(b, a)$
- (M3) $M(\lambda a, \lambda b) = \lambda M(a, b)$

for all $a, b \in \mathbb{R}_+$.

The property (M1) is called internality while (M2) and (M3) are called symmetry and homogeneity, respectively. The condition (M1) is the absolutely essential part of the definition. The conditions (M2) and (M3) are often unnecessary. However, in this book we will treat them to be equally important in defining a mean. Recall that the mean value theorem of integral calculus says that if $f : [a, b] \rightarrow \mathbb{R}$ is a continuous function, then there exists an intermediate value η in $]a, b[$ depending on a and b such that

$$f(\eta(a, b)) = \frac{1}{b-a} \int_a^b f(x) dx.$$

It is easy to see that $\eta(a, b)$ satisfies the properties (M1) and (M2). Now by selecting an appropriate function f , which ensures the homogeneity condition on η , it is possible to generate various existing means.

To construct the arithmetic mean between two positive numbers a and b consider the function $f(x) = x$. Then using the integral mean value theorem, we obtain

$$\begin{aligned} \eta(a, b) &= f(\eta(a, b)) = \frac{1}{b-a} \int_a^b f(x) dx \\ &= \frac{1}{b-a} \int_a^b x dx = \frac{1}{b-a} \frac{b^2 - a^2}{2} = \frac{a+b}{2}. \end{aligned}$$

Thus, the arithmetic mean $A(a, b)$ of two positive numbers a and b has the following integral representation

$$A(a, b) = \frac{1}{b-a} \int_a^b x dx.$$

Likewise an integral representation of the geometric mean can be obtained as follows. Let $f(x) = \frac{1}{x^2}$ and compute

$$\begin{aligned} \frac{1}{\eta^2(a, b)} &= f(\eta(a, b)) \\ &= \frac{1}{b-a} \int_a^b f(x) dx \\ &= \frac{1}{b-a} \int_a^b \frac{1}{x^2} dx \\ &= \frac{1}{b-a} \left[-\frac{1}{x} \right]_a^b \\ &= \frac{1}{b-a} \left[\frac{1}{a} - \frac{1}{b} \right] \\ &= \frac{1}{ab}. \end{aligned}$$

Thus, we have

$$\eta^2(a, b) = ab$$

or

$$\eta(a, b) = \sqrt{ab}.$$

This shows that the geometric mean, $G(a, b)$, has the following integral representation

$$\frac{1}{G^2(a, b)} = \frac{1}{b-a} \int_a^b \frac{1}{x^2} dx.$$

Selecting $f(x) = \ln x$, we can find an integral representation for the identric mean, $I(a, b)$, between two positive numbers a and b . To see this

compute

$$\begin{aligned}
 \ln \eta(a, b) &= f(\eta(a, b)) \\
 &= \frac{1}{b-a} \int_a^b f(x) dx \\
 &= \frac{1}{b-a} \int_a^b \ln x dx \\
 &= \frac{1}{b-a} [x \ln x - x]_a^b \\
 &= \frac{1}{b-a} [b \ln b - a \ln a - (b-a)] \\
 &= \frac{b \ln b - a \ln a}{b-a} - 1 \\
 &= \ln \left(\frac{b^b}{a^a} \right)^{\frac{1}{b-a}} - \ln e \\
 &= \ln \left(\frac{1}{e} \left[\frac{b^b}{a^a} \right]^{\frac{1}{b-a}} \right).
 \end{aligned}$$

Therefore

$$\eta(a, b) = \frac{1}{e} \left[\frac{b^b}{a^a} \right]^{\frac{1}{b-a}}$$

which is the identric mean. Thus, the identric mean can be represented as follows

$$\ln(I(a, b)) = \frac{1}{b-a} \int_a^b \ln x dx.$$

Note that the identric mean like other means is symmetric and homogeneous of degree one. The identric mean, arithmetic mean and the geometric mean satisfy the following ordering

$$\min\{a, b\} \leq G(a, b) \leq I(a, b) \leq A(a, b) \leq \max\{a, b\}.$$

The logarithmic mean $L(a, b)$ can be obtained by selecting an appropriate function $f(x)$ and then applying the integral mean value theorem to it.

If we take $f(x) = \frac{1}{x}$ and apply the mean value theorem, then we have

$$\begin{aligned}\frac{1}{\eta(a,b)} &= f(\eta(a,b)) \\ &= \frac{1}{b-a} \int_a^b f(x) dx \\ &= \frac{1}{b-a} \int_a^b \frac{1}{x} dx \\ &= \frac{1}{b-a} [\ln x]_a^b \\ &= \frac{\ln b - \ln a}{b-a}.\end{aligned}$$

Hence

$$\eta(a,b) = \frac{b-a}{\ln b - \ln a}.$$

Therefore, the integral representation of the logarithmic mean, $L(a,b)$, is

$$\frac{1}{L(a,b)} = \frac{1}{b-a} \int_a^b \frac{1}{x} dx.$$

These integral representations suggest that a further generalization of the above means by applying Theorem 7.2, that is the generalized mean value theorem. Let $g : [a,b] \rightarrow \mathbb{R}$ be a strictly positive, integrable function. By the Theorem 7.2, we have the following generalized means

$$\begin{aligned}\ln(I_g(a,b)) &= \frac{\int_a^b g(x) \ln x dx}{\int_a^b g(x) dx} \\ A_g(a,b) &= \frac{\int_a^b g(x) x dx}{\int_a^b g(x) dx} \\ \frac{1}{L_g(a,b)} &= \frac{\int_a^b g(x) \frac{1}{x} dx}{\int_a^b g(x) dx} \\ \frac{1}{G_g^2(a,b)} &= \frac{\int_a^b g(x) \frac{1}{x^2} dx}{\int_a^b g(x) dx}.\end{aligned}$$

If $g(x) \equiv 1$, then we obtain our familiar means.

We have constructed a large class of means using the generalized integral mean value theorem. The classical means such as arithmetic mean, geometric mean, harmonic mean have found many applications in various branches of mathematics. In the remaining portion of this section, we briefly describe the arithmetic-geometric mean iteration of Gauss (1777-1855) and its connection to classical analysis.

One of the jewels of classical analysis is the arithmetic-geometric mean iteration of Gauss. For two positive numbers a and b let us denote

$$a_0 = a \quad \text{and} \quad b_0 = b$$

and we define recursively

$$a_{n+1} = A(a_n, b_n) = \frac{a_n + b_n}{2}$$

$$b_{n+1} = G(a_n, b_n) = \sqrt{a_n b_n}.$$

If we assume $0 < a < b$, then from the arithmetic-geometric mean inequality, we have

$$a_n \geq a_{n+1} \geq b_{n+1} \geq b_n.$$

Therefore

$$0 \leq b_{n+1} - a_{n+1} = \frac{(b_n - a_n)^2}{2(\sqrt{a_n} + \sqrt{b_n})^2}.$$

Hence

$$\lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} a_n.$$

Let us denote this common limit by

$$M(a, b) = \lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} a_n.$$

In 1791, Gauss first considered the arithmetic-geometric mean (AGM) iteration at the age of 14 and this led him to the discovery of elliptic and modular functions. The limit $M(a, b)$ has many interesting properties. We outline some of these properties without a proof.

The limit $M(a, b)$ is a homogeneous function of degree one, that is

$$M(\lambda a, \lambda b) = \lambda M(a, b).$$

This limit $M(a, b)$ satisfies the functional equation

$$M(a, b) = M\left(\frac{a+b}{2}, \sqrt{ab}\right). \quad (7.15)$$

If we define $f(x) = M(1, x)$, then f satisfies

$$f(x) = \frac{1+x}{2} f\left(\frac{2\sqrt{x}}{1+x}\right).$$

The function $M(1, x)$ can be expressed in closed form in terms of the following complete elliptic integral

$$\frac{1}{M(1, x)} = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1 - (1-x^2)\sin^2 \theta}}.$$

Gauss derived this closed form in 1866. The limit $M(1, x)$ satisfies an hypergeometric differential equation

$$(x^3 - x) \frac{d^2y}{dx^2} + (3x^2 - 1) \frac{dy}{dx} + xy = 0.$$

An integral of the form $\int [f(x) + g(x) \sqrt{(x-a)(x-b)(x-c)}] dx$, where $f(x), g(x)$ are rational functions and a, b, c are distinct constants, can not be integrated in terms of "known" functions. These integrals are called the elliptic integrals since they first occurred in trying to find the arc length of an ellipse. Euler and others studied elliptic integrals. An integral of the form $\int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}}$ is called a complete elliptic integral. Here a and b are real constants. Legendre used the arithmetic-geometric mean iteration to evaluate the complete elliptic integrals. A complete elliptic integral $\int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}}$ is equal to $\frac{\pi}{2M(a, b)}$, that is

$$\int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}} = \frac{\pi}{2M(a, b)}.$$

Thus, the above elliptic integral can be evaluated by evaluating $M(a, b)$.

Modifications of the arithmetic-geometric iteration of Gauss also have been examined by many authors. For an account of various modifications the interested reader should refer to the book of Bullen, Mitrinovic and

Vasic (1988). Here, we present two such modifications. For two positive numbers a and b let us denote

$$a_0 = a \quad \text{and} \quad b_0 = b$$

and if we define recursively

$$a_{n+1} = A(a_n, b_n) = \frac{a_n + b_n}{2}$$

$$b_{n+1} = G(a_{n+1}, b_n) = \sqrt{a_{n+1} b_n},$$

then the sequences $\{a_n\}$ and $\{b_n\}$ have a common limit and it is given by

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = \frac{\sqrt{b^2 - a^2}}{\cos^{-1}(\frac{a}{b})}.$$

Similarly, if we define

$$b_{n+1} = G(a_n, b_n) = \sqrt{a_n b_n}$$

$$a_{n+1} = A(a_n, b_{n+1}) = \frac{a_n + b_{n+1}}{2},$$

then the common limit of these sequences is

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = \frac{\sqrt{a(b-a)}}{\cos(\sqrt{\frac{a}{b}})}.$$

7.3 Coincidence of Mean Values

In this section, we characterize all functions that have the same mean values as their derivative. To prove a characterization theorem, the following result is needed. This result is called the Bernstein Theorem and we refer the reader to the book of Walter (1985) for its proof.

Theorem 7.6 *Let g be infinitely often differentiable on the interval $I =] -r, r[$ and be such that for all large n either $D^n g(x) \geq 0$ for all $x \in I$ or $(-1)^n D^n g(x) \geq 0$ for all $x \in I$, then g is uniquely determined by its Taylor series expansion on the interval I . (Here $D^n g(x)$ denotes the n^{th} derivative of the function g .)*

For the next theorem we will need to impose the following conditions on the function $f : \mathbb{R} \rightarrow \mathbb{R}$:

- (i) f is twice continuously differentiable,
- (ii) $\text{sign}(f'(x)) = c_1$ for all x , where $c_1 \in \{-1, 1\}$,
- (iii) $\text{sign}(f''(x)) = c_2$ for all x , where $c_2 \in \{-1, 1\}$.

The mean value theorems for differential and integral calculus then state that the mean values $W(x, a)$ and $w(x, a)$ are well defined for all $x, a \in \mathbb{R}$ and are given by the following equations for $x \neq a$:

$$f(W(x, a)) = \frac{\int_a^x f(s) ds}{x - a} \quad (7.16)$$

$$f'(w(x, a)) = \frac{f(x) - f(a)}{x - a} \quad (7.17)$$

and for $x = a$ we define $W(a, a) = a = w(a, a)$. In general the two mean values are not the same. Thus the question arises, for which functions f are the mean values $W(x, a)$ and $w(x, a)$ the same for all x and a . The following theorem due to Kranz and Thews (1991) answers this question.

Theorem 7.7 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ satisfy (i), (ii) and (iii) above. If for every $x, a \in \mathbb{R}$ we have that*

$$W(x, a) = w(x, a), \quad (7.18)$$

then there are $\alpha, \beta \neq 0$ and $\mu \in \mathbb{R}$ such that

$$f(x) = \alpha e^{\beta x} + \mu.$$

Proof: We will first show that f is infinitely often differentiable. Let a be fixed. For simplicity we will write $w(x)$ for $w(x, a)$. We thus have

$$\frac{\int_a^x f(s) ds}{(x - a)} = \int_0^1 f(a + t(x - a)) dt. \quad (7.19)$$

Using (7.16) and (7.18), we thus have

$$w(x) = f^{-1} \left(\int_0^1 f(a + t(x - a)) dt \right), \quad (7.20)$$

from which it is easily seen that w is as often differentiable as f is. Differentiating (7.16) with respect to x yields

$$w'(x) = \frac{\int_0^1 f'(a + t(x-a))t dt}{f'(w(x))} > 0.$$

Hence $w^{-1}(y)$ exists and is as often differentiable as f is. We now let $x > a$. With $y = w(x) > a$, it follows from (7.17) that

$$f'(y) = \frac{f(w^{-1}(y)) - f(a)}{w^{-1}(y) - a}.$$

Thus for $x > a$, f' is as often differentiable as f is, whence f is infinitely often differentiable. Since a was arbitrary, this is true everywhere and for w as well.

Now we will show that the k^{th} derivatives of f at a , denoted $D^k f(a)$ are uniquely determined by $f(a), f'(a), f''(a)$, moreover that

$$D^k f(a) = D^k g(a) \text{ for } k = 0, 1, \dots,$$

where

$$\begin{aligned} g(x) &= Ae^{Bx} + C \quad \text{with} \quad B = \frac{f''(a)}{f'(a)}, \\ A &= \frac{f'(a)}{Be^{Ba}} \quad \text{and} \quad C = f(a) - Ae^{Ba}. \end{aligned}$$

This is easily seen as follows: A simple calculation shows that g indeed does satisfy the conditions (7.16)–(7.18) and that $D^k f(a) = D^k g(a)$ for $k = 0, 1, 2$.

It remains to show that we have uniqueness: From (7.16)–(7.18), we get

$$f(w(x)) = \int_0^1 f(a + t(x-a)) dt \tag{7.21}$$

$$f'(w(x)) = \int_0^1 f'(a + t(x-a)) dt \tag{7.22}$$

and differentiating (7.19) and (7.20) n times yields for $x = a$:

$$D^n(f \circ w)(a) = \int_0^1 D^n F(a)t^n dt = \frac{D^n f(a)}{n+1} \tag{7.23}$$

$$D^n(f' \circ w)(a) = \int_0^1 D^{n+1}f(a)t^n dt = \frac{D^{n+1}f(a)}{n+1}. \quad (7.24)$$

For $n = 1$, it follows from (7.22) and (7.23) that

$$\begin{aligned} ((D^n f) \circ w)(a)(w'(a))^n + S_n(a) + ((D^1 f) \circ w)(a)D^n w(a) \\ = \frac{D^n f(a)}{n+1} \end{aligned} \quad (7.25)$$

$$\begin{aligned} ((D^{n+1} f) \circ w)(a)(w'(a))^n + R_n(a) + ((D^2 f) \circ w)(a)D^n w(a) \\ = \frac{D^{n+1} f(a)}{n+1}. \end{aligned} \quad (7.26)$$

Where S_n contains derivatives of f and w of order at most $n - 1$ and R_n contains derivatives of f of order at most n and of w of order at most $n - 1$. Since $f'(a) \neq 0$ and $f''(a) \neq 0$, we can solve (7.24) and (7.25) for $D^n w(a)$. This yields

$$\begin{aligned} \frac{\frac{D^n f(a)}{n+1} - ((D^n f) \circ w)(a)(w'(a))^n - S_n(a)}{D^1 f(a)} \\ = \frac{\left(\frac{1}{n+1} - \left(\frac{1}{2}\right)^n\right)D^{n+1} f(a) - R_n(a)}{D^2 f(a)}. \end{aligned} \quad (7.27)$$

For $n \geq 2$, we can solve (7.26) for $D^{n+1} f(a)$, thus this derivative is determined by the derivatives of lower order (and one also notices that by (7.24) the k^{th} derivative of w is a function of the first k derivatives of f).

Thus all derivatives are determined through the function value and the first two derivatives. It remains to show that f is in fact analytic, since then the theorem of identity yields the result. This immediately follows from the fact that the sign of $B = B(a)$ and $A = A(a)$ are constant and the Bernstein Theorem.

It is clear that all derivatives of f (after possible multiplication by -1) are positive or alternating in sign. This completes the proof of the theorem.

7.4 Some Open Problems

In section 3, we briefly discussed the iteration of arithmetic and geometric means. Readers who wish to know more about this subject should refer to

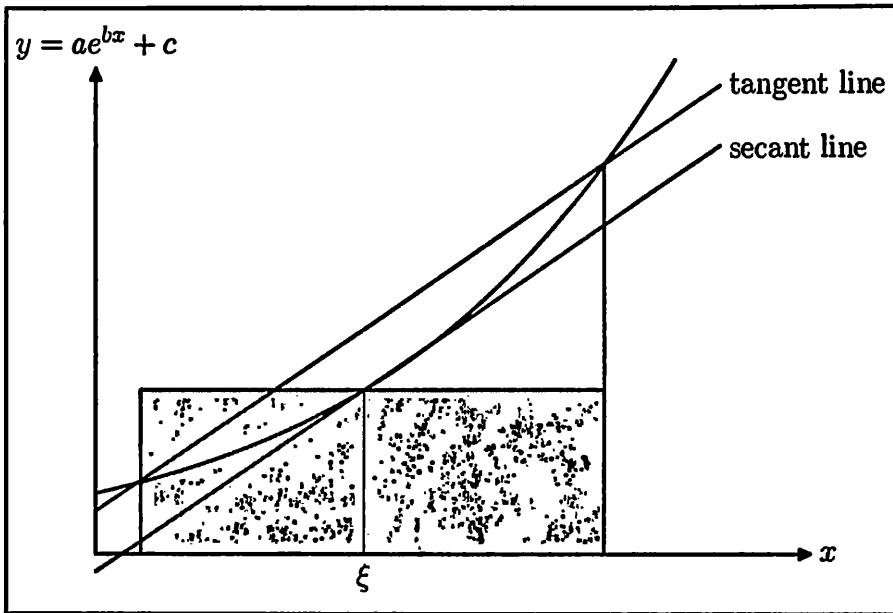


Fig. 7.4 A Geometrical Interpretation of Kranz and Thews Result.

the excellent book by Borwein and Borwein (1987). The Borwein brothers posed some open problems in 1992 regarding the iteration of means. In this section, we present some of their open problems.

Let a and b be two positive real numbers. A *Gaussian mean iteration* associated with two means U and V is the two-term iterations

$$\begin{aligned} a_{n+1} &= U(a_n, b_n) \\ b_{n+1} &= V(a_n, b_n) \end{aligned}$$

with initial values $a_0 = a$ and $b_0 = b$. The common limit of $\{a_n\}$ and $\{b_n\}$ when it exists, is called the *Gaussian compound* of U and V and denoted by

$$U \otimes V = U \otimes V(a, b).$$

If $U(a, b) = A(a, b)$ and $V(a, b) = H(a, b)$, then in closed form $U \otimes V$ is given by

$$U \otimes V(a, b) = \sqrt{ab} = G(a, b).$$

If $U(a, b) = H_p(a, b)$ and $V(a, b) = H_{-p}(a, b)$, then in closed form $U \otimes V$

is given by

$$U \otimes V(a, b) = \sqrt{ab} = G(a, b),$$

where

$$H_p(a, b) = \sqrt[p]{\frac{a^p + b^p}{2}}, \quad p \in \mathbb{R} \setminus \{0\}.$$

If $U(a, b) = A(a, b)$ and $V(a, b) = G(a, b)$, then in closed form $U \otimes V$ is given by

$$U \otimes V(a, b) = M(a, b).$$

Note that we have already discussed some properties of $M(a, b)$.

Let

$$Q(a, b) = \sqrt{\frac{a^2 + b^2}{2}}$$

be the quadratic mean and

$$L_2(a, b) = \frac{a^2 + b^2}{a + b}$$

be a Lehmer mean. One of the open problems of Borwein and Borwein (1992) was to identify, in closed form, the Gaussian compound associated with the arithmetic mean $A(a, b)$ and the quadratic mean $Q(a, b)$. Similarly, the closed form of the Gaussian compound associated with arithmetic mean $A(a, b)$ and the Lehmer mean $L_2(a, b)$ is also not known. It will be nice to know the closed form of these two limits.

The definition of Gaussian compound can be extended to higher dimensions by taking U and V to be two means of n positive real numbers. Given three positive numbers a, b and c , let us denote

$$a_o = a, \quad b_o = b \quad \text{and} \quad c_o = c$$

and we define recursively

$$\begin{aligned} a_{n+1} &= A(a_n, b_n, c_n) = \frac{a_n + b_n + c_n}{3} \\ b_{n+1} &= U(a_n, b_n, c_n) = \frac{a_n b_n + a_n c_n + b_n c_n}{a_n + b_n + c_n} \\ c_{n+1} &= H(a_n, b_n, c_n) = \frac{3a_n b_n c_n}{a_n b_n + a_n c_n + b_n c_n}. \end{aligned}$$

Then Stieltjes (1891) showed that the limit function is $G(a, b, c)$, the geometric mean of a , b and c .

If, on the other hand we define recursively

$$\begin{aligned} a_{n+1} &= A(a_n, b_n, c_n) = \frac{a_n + b_n + c_n}{3} \\ b_{n+1} &= L_2(a_n, b_n, c_n) = \frac{a_n^2 + b_n^2 + c_n^2}{a_n + b_n + c_n} \\ c_{n+1} &= H_2(a_n, b_n, c_n) = \left(\frac{a_n^2 + b_n^2 + c_n^2}{3} \right)^{\frac{1}{2}}, \end{aligned}$$

then the three sequences $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ also converge to a common limit. What can be said about the limit function? In general, what can be said about multidimensional Gaussian compounds?

Recall that the Stolarsky mean $\eta_\alpha(x, y)$ is defined as

$$\eta_\alpha(x, y) = \left(\frac{x^\alpha - y^\alpha}{\alpha(x - y)} \right)^{\frac{1}{\alpha-1}} \quad \text{for all } \alpha \neq 0, 1.$$

It was conjectured by Alzer (1986) that

$$L(x, y) < \frac{\eta_\alpha(x, y) + \eta_{-\alpha}(x, y)}{2} < A(x, y)$$

for all $\alpha \in \mathbb{R} \setminus \{0\}$. Up to now neither a proof nor a counter example is known for this conjecture.

Finally, we conclude this section with the following. The functional equation (7.15) is an interesting functional equation. A variant of (7.15) is the following functional equation:

$$f\left(\frac{a+b}{2}, \frac{2ab}{a+b}\right) = f(a, b), \quad (7.28)$$

where $f : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$. Haruki and Rassias (1995) have studied this functional equation. They have shown that if f can be represented by

$$f(a, b) = \frac{1}{2\pi} \int_0^{2\pi} g(s) d\theta,$$

where $s = a \sin^2 \theta + b \cos^2 \theta$, $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ is a function such that $g''(x)$ is continuous in \mathbb{R}_+ , then the only solution of (7.28) is given by

$$f(a, b) = \frac{A}{\sqrt{ab}} + B,$$

where A and B are arbitrary constants. Haruki and Rassias (1995) posed the following problem. Let $f : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$ be a continuous function in $\mathbb{R}_+ \times \mathbb{R}_+$. Is the only continuous solution of the functional equation (7.28) given by $f(a, b) = \phi(ab)$, where $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}$ is an arbitrary continuous function?

Bibliography

- [1] A. Abian. An ultimate proof of Rolle's theorem. *Amer. Math. Monthly*, 86:485–486, 1979.
- [2] J. Aczél. The general solution of two functional equations by reduction to functions additive in two variables and with aid of Hamel-bases. *Glasnik Mat.-Fiz. Astronom. Drustvo Mat. Fiz. Hrvatske*, 20:65–73, 1965.
- [3] J. Aczél. *Lectures on Functional Equations and Their Applications*. Academic Press, New York-London, 1966.
- [4] J. Aczél. A mean value property of the derivative of quadratic polynomials - without mean values and derivatives. *Math. Magazine*, 58:42–45, 1985.
- [5] J. Aczél. *A Short Course on Functional Equations*. D. Reidel Publishing Company, Dordrecht, Holland, 1987.
- [6] J. Aczél and J. Dhombres. *Functional Equations in Several Variables*. Cambridge University Press, Cambridge, 1989.
- [7] J. Aczél and M. Kuczma. On two mean value properties and functional equations associated with them. *Aequationes Math.*, 38:216–235, 1989.
- [8] H. Alzer. Über Mittelwerte, die zwischen dem geometrischen und dem logarithmischen Mittle zweier Zahlen liegen. *Anz. Österr. Akad. Wiss., Math. -Naturw.*, 123:5–9, 1986.
- [9] A.M. Ampère. Recherche sur quelques points de la theorie des functions. *J. Ecole Polyt.*, Cah., 6:148–181, 1806.
- [10] C.F. Aull. The first symmetric derivative. *Amer. Math. Monthly*, 74:708–711, 1967.
- [11] G. Aumann. Über einen elementararen Zusammenhang zwischen Mittelwerten, Streckenrechnung und Kegelschnitten. *Tohoku Math. J.*, 42:32–37, 1936.

- [12] G. Aumann. Volkommene Funktionalmittel und gewisse Kegelschnitteigenschaften. *J. Für Math.*, 176:49–55, 1936.
- [13] G. Aumann and O. Haupt. *Einführung in die reelle Analysis, Band II*. W. de Gruyter, Berlin, 1979.
- [14] D. Azagra, J. Gomez, and J.A. Jaramillo. Rolle's theorem and negligibility of points in infinite dimensional Banach spaces. *J. Math. Anal. Appl.*, 213:487–495, 1997.
- [15] A.K. Aziz and J.B. Diaz. On Pompeiu's proof of the mean value theorem. *Contributions to Differential Equations*, 1:467–481, 1963.
- [16] D.F. Bailey. A mean-value property of cubic polynomials-without mean value. *Mathematics Magazine*, 65:123–124, 1992.
- [17] D.F. Bailey and G.F. Fix. A generalization of mean value theorem. *Appl. Math. Lett.*, 1:327–330, 1988.
- [18] Z. Bao-lin. A note on the mean value theorem for integrals. *Amer. Math. Monthly*, 104:561–562, 1997.
- [19] L.C. Barret and R.A. Jacobson. Extended laws of the mean. *Amer. Math. Monthly*, 67:1005–1007, 1960.
- [20] L. Bers. On avoiding the mean value theorem. *Amer. Math. Monthly*, 74:583, 1967.
- [21] R.P. Boas. *A Primer of Real Functions*. The Mathematical Association of America, USA, 1981.
- [22] T. Boggio. Sur une proposition de M. Pompeiu. *Mathematica (Cluj)*, 23:101–102, 1947.
- [23] T. Boggio. Sur la formule des accroisements finis. *Acad. Serbe. Sci. Publ. Inst. Math.*, 3:219–226, 1948.
- [24] O. Bonnet. Remarques sur quelques intégrales définies. *Journal de Mathématiques Pure et Appliquées*, 14:249–256, 1849.
- [25] J.M. Borwein and P.B. Borwein. *Pi and the AGM: A study in Analytic Number Theory and Computational Complexity*. Wiley, New York, 1986.
- [26] J.M. Borwein and P.B. Borwein. The way of all means. *Amer. Math. Monthly*, 99:519–522, 1992.
- [27] K.N. Boyadzhiev. Mean value theorems for traces. *Math. Japonica*, 38:217–224, 1993.
- [28] D.M. Bressoud. *A Radical Approach to Real Analysis*. Mathematical Association of America, Washington, DC, 1994.
- [29] R. Brown, T.C. Craven, and M.J. Pelling. Ordered fields satisfying Rolle's theorem. *Illinois J. Math.*, 30:66–78, 1986.

- [30] A. Bruckner. *Differentiation of Real Functions*. American Mathematical Society, Providence, 1994.
- [31] P.S. Bullen, D.S. Mitrinovic, and P.M. Vasic. *Means and Their Inequalities*. D. Reidel Publishing Company, Dordrechat, 1988.
- [32] B.C. Carlson. The logarithmic mean. *Amer. Math. Monthly*, 79:615–618, 1972.
- [33] E. Castagnoli. Estensioni di alcuni classici teoremi dell’analisi al caso di funzioni non derivabili. *Ist. Mat. Fin. Univ. di Parma, serie II*, 11:0–0, 1983.
- [34] A.L. Cauchy. *Cours d’analyse de l’Ecole Polytechnique, Vol 1. Analyse algebrique*. Debure, Paris, 1821.
- [35] B. Choczewski and M. Kuczma. Sur certaines equationes fonctionnelles considerees par I. Stamate. *Mathematica (Cluj)*, 27:225–233, 1962.
- [36] J.K. Chung and P.K. Sahoo. Characterizations of polynomials with mean value properties. *Appl. Math. Lett.*, 6:97–100, 1993.
- [37] F.H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, New-York, 1983.
- [38] F.H. Clarke and Yu.S. Ledyaev. New finite-increment formulas. *Russian Acad. Sci. Dokl. Math.*, 48:75–79, 1993.
- [39] F.H. Clarke and Yu.S. Ledyaev. Mean value inequalities. *Proceedings of the AMS*, 122:1075–1083, 1994.
- [40] L. Cohen. On being mean to the mean value theorem. *Amer. Math. Monthly*, 74:581–582, 1967.
- [41] R. Courant. *Differential and Integral Calculus, Volume II*. Interscience Publishers, New York, 1964.
- [42] G.E. Cross and PL. Kannappan. A functional identity characterizing polynomials. *Aequationes Mathematicae*, 34:147–152, 1993.
- [43] B. Crstici and M. Neagu. About a functional characterization of polynomials. *Inst. Polit. “Traian Vuia” Timisoara. Luc x. Sem. Mat. Fiz. mainov*, 1:59–62, 1987.
- [44] G. Darboux. Memoire sur les fonctions discontinues. *Ann. Sci. Scuola. Sup.*, 4:57–112, 1875.
- [45] Z. Daroczy and A. Jarai. On the measurable solution of a functional equation of the information theory. *Acta Math. Acad Sci. Hungaricae*, 34:105–116, 1979.
- [46] R.O. Davies and G. Rousseau. A divided-difference characterization of polynomials over a general field. *Aequationes Math.*, 55:73–78, 1998.

- [47] R.M. Davitt, R.C. Powers, T. Riedel, and P.K. Sahoo. Flett's mean value theorem for holomorphic functions. *Math. Magazine* (To appear), 1998.
- [48] J.B. Diaz and R. Vyborny. On a mean value theorem of the weak differential calculus of vector valued functions. *Contributions to Differential Equations*, 1:271–273, 1963.
- [49] J.B. Diaz and R. Vyborny. On mean value theorems for strongly continuous vector valued functions. *Contributions to Differential Equations*, 3:107–118, 1964.
- [50] J.B. Diaz and R. Vyborny. A fractional mean value theorem, and Taylor theorem, for strongly continuous vector valued functions. *Czechoslovak Math. J.*, 15:299–304, 1965.
- [51] J.B. Diaz and R. Vyborny. On some mean value theorems of the differential calculus. *Bull. Austral. Math. Soc.*, 5:227–238, 1971.
- [52] J. Dieudonne. Sur une generalisation du theoreme de Rolle aux fonctions d'une variable complexe. *Ann. of Math.*, 31:79–116, 1930.
- [53] E3338. A characterizatic of low degree polynomials. *Amer. Math. Monthly*, 98:268–269, 1991.
- [54] R.J. Easton and S.G. Wayment. A mean value theorem. *Amer. Math. Monthly*, 77:170–172, 1970.
- [55] B.R. Ebanks, P.K. Sahoo, and W. Sander. Determination of measurable sum form information measures satisfying $(2, 2)$ -additivity of degree (α, β) . *Radovi Matematicki*, 6:77–96, 1990.
- [56] C.J. Eliezer. A solution to $f\{(1-x)(1-y)\} + f(xy) = f\{x(1-y)\} + f\{y(1-x)\}$. *Aequationes Math.*, 10:311–312, 1974.
- [57] J.P. Evans. The extended law of the mean by a translation-rotation of axes. *Amer. Math. Monthly*, 67:580–581, 1960.
- [58] J.C. Evard and F. Jafari. A complex Rolle's theorem. *Amer. Math. Monthly*, 99:858–861, 1992.
- [59] J. Ferrer. Rolle's theorem fails in l_2 . *Amer. Math. Monthly*, 103:161–165, 1996.
- [60] T.M. Flett. A mean value theorem. *Math. Gazette*, 42:38–39, 1958.
- [61] T.M. Flett. Some historical notes and speculations concerning the Mean Value Theorems of differential calculus. *Bull. Inst. Math. Appl.*, 10:66–72, 1974.
- [62] M. Furi and M. Martelli. On the mean value theorem, inequality, and inclusion. *Amer. Math. Monthly*, 98:840–847, 1991.
- [63] M. Furi and M. Martelli. A multidimensional version of Rolle's theorem. *Amer. Math. Monthly*, 102:243–249, 1995.

- [64] L. Gajek and D. Zagrodny. Geometric mean value theorems for Dini derivatives. *J. Math. Anal. Appl.*, 191:56–76, 1995.
- [65] C.F. Gauss. *Theoria motus corporum caelestium*. Perthes-Besser, Hamburg, 1809.
- [66] G. Giorgi and S. Komlosi. Dini derivatives in optimization - Part I. *Revista di Matematica per le scienze economiche e sociali*, 10:3–30, 1992.
- [67] D.B. Goodner. Mean value theorems for functions with finite derivatives. *Amer. Math. Monthly*, 67:852–855, 1960.
- [68] G. Hamel. Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung $f(x + y) = f(x) + f(y)$. *Math. Ann.*, 60:459–462, 1905.
- [69] H. Haruki. New characterizations of the arithmetic-geometric mean of Gauss and other well-known mean values. *Publ. Math. Debrecen*, 38:323–332, 1991.
- [70] H. Haruki and Th. M. Rassias. A new analogue of Gauss functional equation. *Internat. J. Math. & Math. Sci.*, 18:749–756, 1995.
- [71] Sh. Haruki. A property of quadratic polynomials. *Amer. Math. Monthly*, 86:577–579, 1979.
- [72] J.B. Hiriart-Urruty. Mean value theorems in nonsmooth analysis. *Numer. Funct. Anal. Optimization*, 2:1–30, 1980.
- [73] R.A. Horn. Editor's endnotes . *Amer. Math. Monthly*, 105:494–494, 1998.
- [74] A. Horwitz. Means and divided differences. *Jour. Math. Anal. Appl.*, 191:618–632, 1995.
- [75] E. Isaacson and H.B. Keller. *Analysis of Numerical Methods*. Wiley, New York, 1966.
- [76] B. Jacobson. On the mean value theorem for integrals. *Amer. Math. Monthly*, 89:300–301, 1982.
- [77] G.J. Janusz. *Calculus*. Wm. C. Brown, Dubuque, Iowa, 1994.
- [78] E. Kamke. Zur definition der affinen abbildung. *Jahresbericht der Deutschen Mathematischer-Vereinigung*, 36:145–156, 1927.
- [79] PL. Kannappan. Divided differences and polynomials. *C.R. Math. Report, Canada*, 16:187–192, 1994.
- [80] PL. Kannappan and B. Crstici. Two functional identities characterizing polynomials. Itinerant Seminar on Functional Equations, Approximation and Convexity, Cluj-Napoca, 175–180, 1989.
- [81] PL. Kannappan and P.G. Laird. Functional equations and second degree polynomials. *Jour. Ramanujan Math. Soc.*, 5:95–110, 1993.

- [82] PL. Kannappan, T. Riedel, and P.K. Sahoo. On a generalization of a functional equation associated with Simpson's rule. *Results in Mathematics*, 31:115–126, 1997.
- [83] PL. Kannappan, T. Riedel, and P.K. Sahoo. On a functional equation associated with Simpson's rule. *Wyz. Szkoła Ped. Krakow Rocznik Nauk.-Dydakt. Prace Matematyczne* (To appear), 1998.
- [84] PL. Kannappan and P.K. Sahoo. A property of quadratic polynomials in two variables. *Jour. Math. Phy. Sci. (Madras)* (To appear), 1995.
- [85] PL. Kannappan and P.K. Sahoo. Characterization of polynomials and the divided difference. *Proc. Indian Acad. Sci. (Math. Sci.)*, 105:287–290, 1995.
- [86] PL. Kannappan, P.K. Sahoo, and M.S. Jacobson. A characterization of low degree polynomials. *Demonstratio Mathematica*, 28:87–96, 1995.
- [87] L.A. Khan. Mean value theorem in topological vector spaces. *C.R. Math. Rep. Acad. Sci. Canada*, 19:24–27, 1997.
- [88] L. Kranz and K. Thews. Zur Übereinstimmung der Mittelwertstellen von Funktionen und ihren Ableitungen. *El. Math.*, 46:143–145, 1991.
- [89] M. Kuczma. *An Introduction to the Theory of Functional Equations and Inequalities*. Uniwersytet Śląski, Warszawa-Krakow-Katowice, 1985.
- [90] M. Kuczma. On a Stamate-type functional equation. *Publ. Math. Debrecen*, 39:325–338, 1991.
- [91] M. Kuczma. On the quasiarithmetic mean in a mean value property and the associated functional equation. *Aequationes Mathematicae*, 41:33–54, 1991.
- [92] T.V. Lakshminarasimhan. A mean value theorem – an extension. *Amer. Math. Monthly*, 73:862–863, 1966.
- [93] L. Larson. The symmetric derivative. *The Transactions of the AMS*, 277:589–599, 1983.
- [94] E.B. Leath and M.C. Sholander. Extended mean values. *Amer. Math. Monthly*, 85:84–90, 1978.
- [95] E.B. Leath and M.C. Sholander. Extended mean values II. *J. Math. Anal. Appl.*, 92:207–223, 1992.
- [96] E.B. Leath and M.C. Sholander. Minkowski's inequality for two variable difference means. *Proc. Amer. Math. Soc.*, 126:779–789, 1998.
- [97] A.M. Legendre. *Elements de geometrie. Note II*. Didot, Paris, 1791.
- [98] Gy. Maksa. Problem. *Aequationes Math.*, 46:301, 1993.
- [99] M. Marsden. The search of Rolle's theorem in the complex domains. *Amer. Math. Monthly*, 92:643–650, 1992.

- [100] M. Mays. Functions which parameterize means. *Amer. Math. Monthly*, 90:677–683, 1983.
- [101] R. McLeod. Mean value theorems for vector-valued functions. *Proc. Edin. Math. Soc.*, 14:197–209, 1965.
- [102] A.M. Ostrowski. *Solution of Equations in Euclidean and Banach Spaces*. Academic Press, New York and London, 1973.
- [103] J.P. Penot. On the mean value theorems. *Optimization*, 19:147–156, 1988.
- [104] A.O. Pittenger. The logarithmic mean in n variables. *Amer. Math. Monthly*, 92:99–104, 1985.
- [105] M.J. Poliferno. A natural auxiliary function for the mean value theorem. *Amer. Math. Monthly*, 69:45–47, 1962.
- [106] D. Pompeiu. Sur une proposition analogue au theoreme des accroissements finis. *Mathematica (Cluj)*, 22:143–146, 1946.
- [107] R.C. Powers, T. Riedel, and P.K. Sahoo. Limit properties of differential mean values. *J. Math. Appl. Anal.* (To appear), 1998.
- [108] G. Prasad. *Six Lectures on the Mean-Value Theorem of the Differential Calculus*. University of Calcutta, Calcutta, India, 1931.
- [109] J. Rätz and D. Russell. An extremal problem related to probability. *Aequationes Math.*, 34:316–324, 1987.
- [110] S. Reich. On mean value theorems. *Amer. Math. Monthly*, 76:70–73, 1969.
- [111] T. Riedel and P.K. Sahoo. A functional equation characterizing low degree polynomials in two variables. *Demonstratio Mathematica*, 30:85–94, 1997.
- [112] J.M. Robertson. A local mean value theorem for the complex plane. *Proc. Edin. Math. Soc.*, 16:329–331, 1968.
- [113] I. Rosenholtz. A topological mean value theorem for the plane. *Amer. Math. Monthly*, 98:149–153, 1991.
- [114] W. Rudin. Problem E3338. *Amer. Math. Monthly*, 96:641, 1989.
- [115] M. Sablik. A remark on a mean value property. *C.R. Math. Rep. Acad. Sci. Canada*, 24:207–212, 1992.
- [116] P.K. Sahoo. Three open problems in functional equations. *Amer. Math. Monthly*, 102:741–742, 1995.
- [117] A. Samuelsson. A local mean value theorem for analytic functions. *Amer. Math. Monthly*, 80:45–46, 1973.
- [118] J. Sander. On the identric and logarithmic means. *Aequationes Math.*, 40:261–270, 1990.
- [119] D.E. Sanderson. A versatile vector mean value theorem. *Amer. Math. Monthly*, 79:381–383, 1972.

- [120] M. Sayrafiezadeh. A generalization of the mean value theorem for integrals. *College Math. Journal*, 26:223–224, 1995.
- [121] I.J. Schoenberg. A conjectured analogue of Rolle’s theorem for polynomials with real or complex coefficients. *Amer. Math. Monthly*, 93:8–13, 1986.
- [122] L. Schumaker. *Spline Functions*. Wiley, New York-Toronto, 1981.
- [123] J. Schwaiger. On a characterization of polynomials by divided differences. *Aequationes Math.*, 48:317–323, 1994.
- [124] H.N. Shapiro. A micronote on a functional equation. *Amer. Math. Monthly*, 80:1041, 1973.
- [125] S.A. Shkarin. On Rolle’s theorem in infinite dimensional Banach spaces. *Mat. Zametki*, 51:128–136, 1992.
- [126] J. Smital. *On Functions and Functional Equations*. Adam Hilger, Bristol and Philadelphia, 1988.
- [127] M.R. Spiegel. Mean value theorems and Taylor series. *Amer. Math. Monthly*, 29:263–266, 1956.
- [128] I. Stamate. On a proposition of Pompeiu (Roumanian). Inst. Politehn. Cluj Lucrari Stiint., 75–78, 1958.
- [129] I. Stamate. A property of the parabola and integration of a functional equation (Roumanian). Inst. Politehn. Cluj Lucrari Stiint., 101–106, 1959.
- [130] D.P. Stanford. Functions satisfying a mean value property at their zeros. *Amer. Math. Monthly*, 80:665–667, 1973.
- [131] K.B. Stolarsky. Generalizations of the logarithmic mean. *Math. Magazine*, 48:87–92, 1975.
- [132] K.B. Stolarsky. The power and generalized logarithmic means. *Amer. Math. Monthly*, 87:545–548, 1980.
- [133] H. Swann. Commentary on rethinking rigor in calculus: The role of the mean value theorem. *Amer. Math. Monthly*, 104:241–245, 1997.
- [134] B.S. Thomson. *Symmetric Properties of Real Functions*. Marcel Dekker, New York-Bassel-Hong Kong, 1994.
- [135] A. Tineo. A generalization of Rolle’s theorem and application to nonlinear equation. *J. Austr. Math. Soc. Ser. A*, 46:395–401, 1989.
- [136] J. Tong and P.A. Braza. A converse of the mean value theorem. *Amer. Math. Monthly*, 104:939–942, 1997.
- [137] D.H. Trahan. A new type of mean value theorem. *Math. Magazine*, 39:264–268, 1956.
- [138] T.W. Tucker. Rethinking rigor in calculus: The role of the mean value theorem. *Amer. Math. Monthly*, 104:231–240, 1997.

- [139] V.B. Uvarov. *Mathematical Analysis*. MIR Publishers, Moscow, 1988.
- [140] W. Walter. *Analysis 1*. Springer, Berlin-Heidelberg-New-York, 1985.
- [141] S.G. Wayment. An integral mean value theorem. *Math. Gazette*, 54:300–301, 1970.
- [142] A. Wilansky. Additive functions. In Kenneth O. May, editor, *Lectures on Calculus*, pages 97–124, Toronto, 1967. Holden-Day, Inc.
- [143] R.C. Yates. The law of the mean. *Amer. Math. Monthly*, 66:583–584, 1959.

Index

- Aczél, J., 1, 7, 40, 51, 57, 85, 103, 109
Alzer, H., 230
Ampère, A.M., 25
Arithmetic-geometric mean
 inequality, 222
 iteration, 222
Aull, C.F., 188, 189, 192
Aumann, G., 78
Automorphism, 18

Bailey, D.F., 51, 53, 55, 56
Banach space, 180
Bao-lin, Z., 64, 76
Barrett, L.C., 149
Bernoulli's inequality, 31
Bernstein Theorem, 224
Boggio, T., 83, 92
Bonnet, O., 25, 214
Borwein, J.M., 228, 229
Borwein, P.B., 228, 229
Bressoud, D.M., 214
Bruckner, A., 206
Bullen, P.S., 224

Castagnoli, E., 201
Cauchy, A.L., 1, 2, 77
Central difference quotient, 182
Chain rule of differentiation, 36
Chung, J.K., 85
Clarke, F.H., 147, 159, 160

Closed interval, 175
Co-ordinate functions, 163
Complete elliptic integral, 223
Complex conjugate, 15, 16
Complex numbers system, 15
Conjecture, 180, 230
Continuous extension, 60
Convex combination, 29
Convex set, 157
Courant, R., 127
Critical point, 160
Crstici, B., 53

Darboux property, 191
Darboux, G., 191
Daroczy, Z., 22, 23
Davitt, R.M., 173
Dense in reals, 8
Determinant, 148
Dieudonne, J., 169
Difference quotient, 182
Differentiable function, 182
Dini derivatives, 195
Divided differences, 40

Ebanks, B.R., 22
Eliezer, C.J., 23
Elliptic integral, 223
Euclidean inner product, 156
Evard, J.C., 147, 169, 173

- Extreme point, 164
- Flett, T.M., 147, 151
- Function
- additive, 1
 - continuous, 2
 - discontinuous, 6
 - almost open, 23
 - analytic, 17
 - biadditive, 18, 131
 - graph, 7
 - holomorphic, 169, 172, 173, 177
 - linear, 2
 - locally increasing, 199
 - locally integrable, 3
 - measurable, 22
 - multiplicative, 12
 - quadratic, 129, 135
 - rationally homogeneous, 4
 - skew-symmetric, 22
- Functional equation, 22, 80, 85, 92, 125, 127, 144, 145, 230
- generalized Jensen, 129
 - Cauchy, 1, 2, 50
 - Jensen, 131
 - Sincov, 97
- Fundamental theorem
- of calculus, 31, 207
- Furi and Martelli's conjecture, 180
- Furi, M., 147, 149, 164, 180
- Gauss, C.F., 2
- Gaussian compound, 228
- Gaussian mean iteration, 228
- Generalized binomial coefficient, 69
- Generalized derivative, 181, 206
- Giorgi, G., 195
- Hölder inequality, 33
- Hamel basis, 9
- Hamel, G., 7
- Haruki, H., 230, 231
- Haruki, Sh., 40, 109
- Hilbert space, 180
- Homogeneity, 218
- Horn, R.A., 28
- Horwitz, A., 62, 63, 81
- Hypergeometric differential equation, 223
- Imaginary part, 15
- Information theory, 32
- Integral representation
- arithmetic mean, 219
 - generalized means, 221
 - geometric mean, 219
 - identric mean, 220
 - logarithmic mean, 38, 221
 - of divided differences, 58
- Intermediate value property, 191
- Intermediate value theorem, 29, 61, 211, 217
- Internality, 218
- Isaacson, E., 58
- Iteration of means, 207
- Jacobian of a function, 163
- Jacobson, B., 64, 69, 76
- Jacobson, M.S., 43, 109, 117
- Jacobson, R.A., 149
- Jafari, F., 147, 169, 173
- Janusz, G.J., 25
- Jarai, A., 22, 23
- Jensen equation, 131
- Jump discontinuity, 186
- Kannapan, PL., 43, 53, 109, 117, 129, 135
- Keller, H.B., 58
- Komlosi, S., 195
- Kranz, L., 225
- Kuczma, M., 1, 57, 83, 85, 86, 91, 92, 125
- Lakshminarasimhan, T.V., 203
- Ledyayev, Yu.S., 147, 159, 160
- Legendre, A.M., 2
- Liminf, 197

- Limiting behavior of means, 64, 69, 70
 Limsup, 197
 Line segment, 175
 Linearly independent, 116
 Local mean value theorem, 177
 Maksa, Gy., 22
 Marsden, M., 169
 Martelli, M., 147, 149, 164, 180
 Mays, M., 62
 McLeod's theorem, 162
 McLeod, R., 147, 162
 Mean, 218
 - arithmetic, 37, 92
 - functional, 61
 - geometric, 37
 - identric, 38
 - Lehmer, 229
 - logarithmic, 37
 - quasiarithmetic, 57
 - Stolarsky, 36, 230
 Mean value theorem
 - a graphical proof, 27
 - geometrical interpretation, 27, 30
 - Cauchy, ix
 - for divided differences
 - Cauchy, 78
 - Lagrange, 58, 60, 69, 78
 - for holomorphic functions, 169
 - Evard and Jafari, 175
 - functions in one variable
 - Cauchy, 77, 166
 - Flett, 151, 152
 - Lagrange, 25, 147
 - Pompeiu, 83
 - functions in one variables
 - Cauchy, 161
 - functions in two variables, 163
 - Cauchy, 144
 - Lagrange, 127, 156
 - Pompeiu, 145
 - integral, 207
 - Bonnet, 214

Flett, 210
 generalization, 208
 Lagrange, vii
 nondifferentiable functions

 - Flett, 203
 - Lagrange, 201
 - Trahan, 204

of integral calculus, 207
 Pompeiu, viii
 symmetrically differentiable functions

 - Cauchy, 194
 - Flett, 193
 - vector valued functions, 160

Mean value type equation, 129

 - generalization, 136

Mitrinovic, D.S., 224
 Multidimensional

 - Flett's theorem, 168
 - Rolle's theorem, 163

Neagu, M., 53
 Nested intervals, 29
 Norm, 156
 Open ball, 164
 Open problem, 22, 80, 125, 145, 228
 Optimization problem, 181
 Ostrowski, A.M., 58
 Partial derivative, 127, 135
 Partial derivatives, 156
 Partition, 212
 Pompeiu, D., 83
 Powers, R.C., 64, 70, 173
 Quadratic polynomial, 135, 145
 Quasi-mean value theorem, 190
 Rätz, J., 78
 Rassias, Th.M., 230, 231
 Rational linear combination, 8
 Real part, 15
 Riedel, T., 64, 70, 135, 173
 Riemann sum, 212

- Rolle's theorem, 26, 166
for symmetrically differentiable
functions, 189
- Rouche's theorem, 177
- Russell, D., 78
- Sahoo, P.K., 22, 43, 53, 64, 70, 85,
109, 117, 129, 135, 173
- Samuelsson, A., 147, 177
- Sander, W., 22
- Sanderson, D.E., 147, 161
- Sayrafiezadeh, M., 212
- Schumaker, L., 80
- Schwaiger, J., 57
- Segment, 157, 159
- Shapiro, H.N., 3
- Simpson's rule, 98
- Sincov equation, 97
- Smital, J., 1
- Stolarsky, K.B., 62, 76
- Swann, H., 27
- Symmetric derivative, 182
- Symmetry, 218
- Taylor's formula, 72
- Taylor's polynomial, 75
- Taylor's theorem, 178
- Thews, K., 225
- Thomson, B.S., 182
- Trahan, D.H., 147, 154, 158, 176, 193,
204
- Tucker, T.W., 28
- Uvarov, V.B., 72
- Vasic, P.M., 224
- Wayment, S.G., 207, 210
- Wilansky, A., 1

ISBN 981-02-3544-5



9 789810 235444

<http://www.worldscientific.com/>

3857 hc