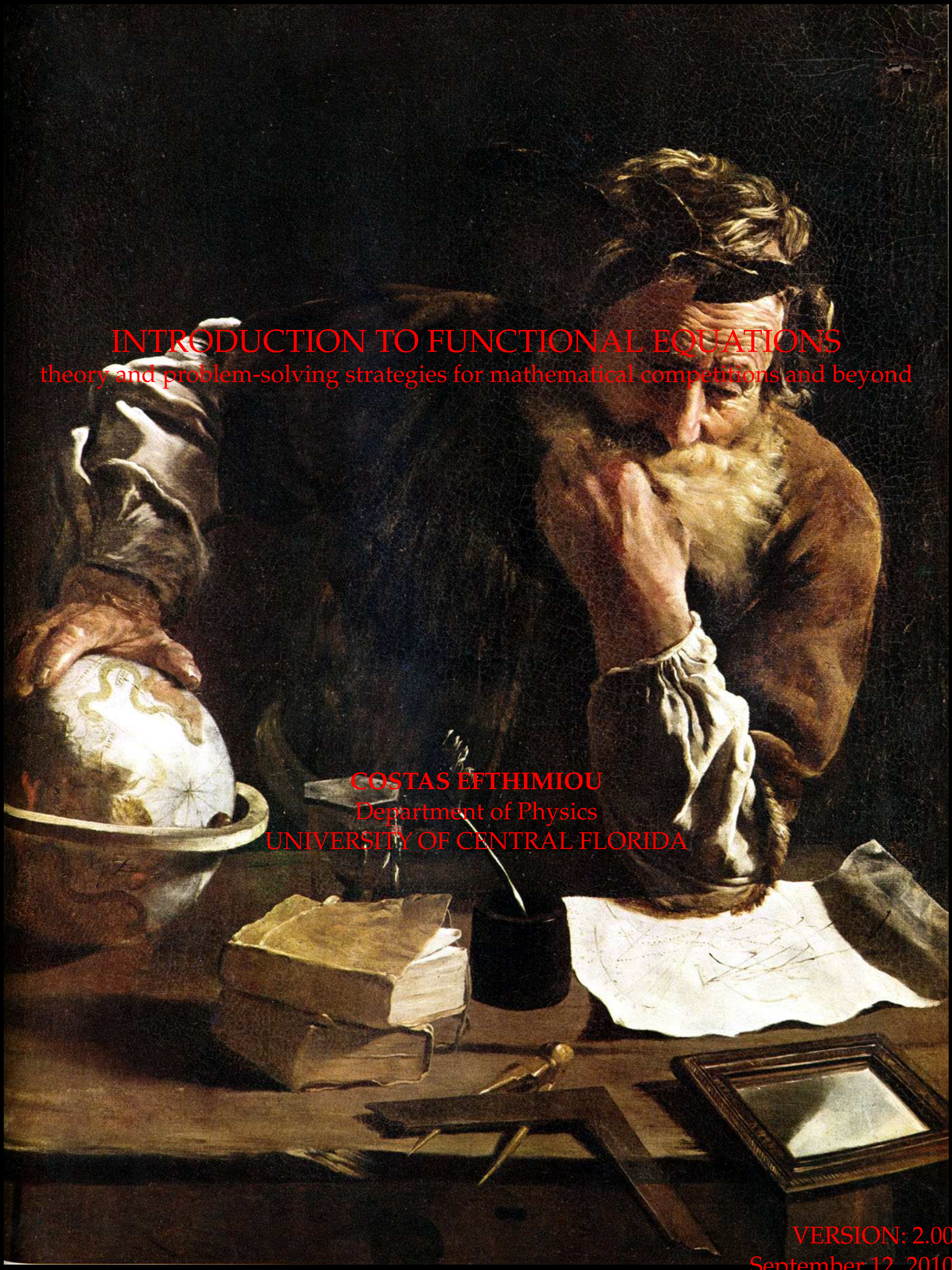

INTRODUCTION TO FUNCTIONAL EQUATIONS

theory and problem-solving strategies for mathematical competitions and beyond

COSTAS EFTHIMIOU
Department of Physics
UNIVERSITY OF CENTRAL FLORIDA

VERSION: 2.00
September 12, 2010

Cover picture: Domenico Fetti's *Archimedes Thoughtful*, Oil on canvas, 1620. State Art Collections Dresden.

♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡

♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡

♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡

♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ **To My Mother**

♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡

♡ who taught me that love is not
measured by how many things
one can offer to a child but
by how many things
a parent sacrifices
for the child

♡

Contents

Introduction

One of the most beautiful mathematical topics I encountered as a student was the topic of functional equations — that is, the topic that deals with the search of functions which satisfy given equations, such as

$$f(x+y) \;=\; f(x)+f(y)\,.$$

This topic is not only remarkable for its beauty but also impressive for the fact that functional equations arise in all areas of mathematics and, even more, science, engineering, and social sciences. They appear at all levels of mathematics. For example, the definition of an even function is a functional equation,

$$f(x) = f(-x)\,,$$

that is encountered as soon as the notion of a function is introduced. At the other extreme, in the forefront of research, during the last two to three decades, the celebrated Young-Baxter functional equation has been at the heart of many different areas of mathematics and theoretical physics such as lattice integrable systems, factorized scattering in quantum field theory, braid and knot theory, and quantum groups to name a few. The Young-Baxter equation is a system of N^6 functional equations for the $(N^2 \times N^2)$-matrix $R(x)$ whose entries are functions of x:

$$\sum_{\alpha,\beta,\gamma=1}^{N} R_{jk}^{\alpha\beta}(x-y)\, R_{i\alpha}^{\ell\gamma}(x)\, R_{\gamma\beta}^{mn}(y) \;=\; \sum_{\alpha,\beta,\gamma=1}^{N} R_{ij}^{\alpha\beta}(y)\, R_{\beta k}^{\gamma n}(x)\, R_{\alpha\gamma}^{\ell m}(x-y)\,,$$

where i, j, k, ℓ, m, n take the values $1, 2, \cdots, N$. Amazingly, although the Yang-Baxter equation appears to impose more constraints than the number of unknowns (N^6 equations for N^4 unknowns), it has a rich set of solutions. The study of this equation is well beyond the goals of the book; however, the reader will hopefully come to appreciate the role that functional equations play in mathematics and science over the course of reading this book.

Despite being such a rich subject, this topic does not fit perfectly into any of the conventional areas of mathematics and thus a systematic way of studying functional equations is not found in any traditional book used in the traditional education of students. Several good books on the topic exist but, unfortunately, they are either relatively hard to find [39] or quite advanced [7, 8, 27]. While this book was in the making, a nice, simple book [38] appeared which has considerable overlap with, and is complementary to, this book.

We should point out in advance that there is no universal technique for solving functional equations, as will become obvious once the reader solves the problems throughout this book. At the current time, thorough knowledge of many areas of mathematics, lots of imagination, and, perhaps, a bit of luck are the best methods for solving such equations.

The core of the book is the result of a series of lectures I presented to the UCF Putnam team after my arrival at UCF. My personal belief is that the training of a math team should be based on a systematic approach for each subject of interest for the mathematical competitions (e.g. functional equations, inequalities, finite and infinite sums, etc.) instead of randomly solving hard problems to get experience. Therefore, this book is an attempt to present the fundamentals of the topic at hand in a pedagogical manner, accessible to students who have a background on the theory of functions up to differentiability. Although there are some parts of the book that use additional ideas from calculus, such parts may be omitted at first reading. In particular, the book should be useful to high school students who participate in math competitions and have an interest in International Math Olympiads (IMO). Many of the problems I have included are problems I solved as a high school student[1] in preparation for the national math exam of my home country, Greece. I still enjoy them as much as I enjoyed them then. In fact, I now appreciate their beauty much more, as I have grown to have a better understanding of the subject and its importance. College students with an interest in the Putnam Competition should also find the book useful, as standard texts do not provide a thorough coverage of functional equations. Finally, we hope that any person with interest in mathematics will find in this book something to his or her interest. However, a piece of advice is in order here for all readers: The majority of the problems discussed in this book are related to mathematical competitions which target ingenuity and insight, so they are not easy. Approach them with caution. They should serve as a source of enjoyment, not despair!

And a final comment: The interested reader should, perhaps, read the book simultaneously with Smìtal's [39] and Small's [38] books, which are at the same level. I cover more functional equations but Smìtal presents beautifully the topic of iterations and functional equations of one variable[2]. Similarly, Small's book [38] is a very enjoyable, well written book and focuses on the most essential aspects of functional equations. Once the reader is done with these three books, he may read Aczél's and Kuczma's authoritative books [7, 8, 27], which contain a vast amount of information. Beyond that level, the reader will be ready to immerse himself in the current literature on the subject, and perhaps even conduct his own research in the area.

More on How to Approach This Book

Mathematics has a reputation of being a very stiff subject that only follows a pre-determined pattern: definitions, axioms, theorems, proofs. Any deviations from this pre-established pattern are neither welcome nor wise. In this respect, mathematics appears to be very different from sciences and, in particular, theoretical physics, which is the closest discipline. For, mathematics is using **demonstrative reasoning** and sciences use **plausible reasoning**.[3]

As was mentioned previously, the majority of the problems in this book are taken from mathematical competitions. As such, these problems are precious not only for their origi-

[1] Yes, I still have many of my notes! As a result, some of the problems I present might be taken from national or local competitions without reference. I would appreciate a note from readers who discover such omissions or other typos and mistakes.

[2] One can detect some influence from Smìtal's book on my presentation in Chapters 16 and 17. I highly recommended this book to any serious student.

[3] The terms are discussed in G. Pólya's classic two-volume work [31] which teaches students the role of guessing in rigorous mathematics and how to become effective guessers.

nality and beauty, but also for the lessons they teach the students regarding mathematical discovery. Without doubt, the final solutions of these problems must be presented as demonstrative reasoning. However, to reach this stage, the students must first go through the stage of plausible reasoning. Solving a problem first involves guessing the solution. Proving the solution, first involves visualization of the proof. Polishing the proof requires trying these steps over and over. Therefore demonstrative reasoning and plausible reasoning are not two distinct and isolated methods; they are just the two faces of the process of discovering new results in mathematics. So, in agreement with Pólya's motto, the problems in this book should be used to "let us learn proving, but also *let us learn guessing*."

Unfortunately, due to lack of time and space, I have omitted the plausible reasoning which can be experienced by attending Math Circles, and I have presented only the demonstrative reasoning in what I believe to be a polished way. However, the students should be assured that the solutions presented could not have been found without some kind of hunches and guesses. The beginning student should thus neither be discouraged nor be disappointed if his reasoning, either at the plausible or at the demonstrative stage, fails to be as good as those of the experienced solvers. Paul Erdös, one of the most prolific publishers of papers in mathematical history and extraordinary problem solver, has said: "Nobody blames a mathematician if the first proof of a new theorem is clumsy." Therefore, solve these problems as many times as necessary to improve your solutions. With every new solution you discover, you gain invaluable knowledge that becomes your new weapon to attack new problems. This is really no new idea and it has been quite familiar to anyone who has tried problem solving. Here is René Descartes' testimony: "Each problem that I solved became a rule which served afterwards to solve other problems." Any reader who places considerable amount of effort on the problems will benefit from them, even if he does not completely solve all the problems.

Of course, all this assume that you see the beauty in these problems. I cannot define mathematical beauty, but I know it when I see it. It is really love at first sight. Upon completion of the first draft of the manuscript, I showed it to some good students ('good' being defined by high GPA in the standard curriculum) who had not attended my lectures. Of course, I was expecting to hear great comments, not so much for the book itself but the collection of the problems. Unfortunately, at least one of them declared these problems 'trick problems'! He appeared unimpressed and, I think, completely uninterested in them. I had always assumed, even after many years of teaching, that if a student stands well above the crowd in mathematical ability, then he also sees well beyond the crowd. However, this incident made me realize that even some students with mathematical talent, when they have been educated in a dry and boring traditional system, often by unmotivated teachers, have never developed mathematical aesthetics and deep understanding of how the process of mathematical discovery really happens. As a blind person, who is otherwise a skillful diamond cutter, looking towards the diamond never sees the sparkling to appreciate the fine cuts, a student who has never be shown 'true mathematics' but is otherwise a skillful solver, can never appreciate the extraordinary beauty of math problems.

Therefore, it is my obligation as I present this book to the students to declare: Trick mathematics is the norm. Straightforward proofs are the exceptions. There are patterns but these are the products of experience and hard work of the giants before us. To unveil a pattern, you must first produce many trick proofs. The trickier the proof, the more the hidden beauty. Such is, for example, Šarkovskii's theorem stated in Chapter 16 — a fine jewel of modern mathematics. Appreciating and solving the problems given in mathematical competitions builds strong intuition, develops mathematical ingenuity, and promotes deep understanding for trick proofs. It leads the way to creating new generations of outstanding mathematicians. Dear reader, please approach this book with this spirit in mind.

Costas Efthimiou
Orlando, FL
July 2010

Acknowledgements

First of all, I must thank the UCF students who sat in my lectures in 2001. Without them, I would never had written the initial notes from which this book materialized.

After the people who inspired the project, the author of any book should thank Donald Knuth, Leslie Lamport and the LaTeX3 Project team for their invaluable services to all authors. With the powerful typesetting language of TeX(and its derivative of LaTeX) any author can easily pen his thoughts in what he thinks is the most effective way. The preparation of this book would have been otherwise a painful or impossible task. The figures in the present book were created using Daniel Luecking's MFPIC package, and I also thank him for his help with the package to produce stand-alone eps figures.

However, from the initial moment of the birth of the project and to the final moment of its completion, many other people helped in various ways. Roy Choudhury and Vaggelis Alexopoulos skimmed over a semi-finished version and gave me few words of approval that made me very happy and reinforced my intention to produce a final copy for publication. Joseph Brennan read the chapter on polynomial and offered some suggestions. Li Zhou contributed a few problems; Caleb Wiese typed a few pages of handwritten notes on difference equations; Robert van Gorder proofread the manuscript and suggested several corrections. Christopher Frye proofread carefully the manuscript near completion and detected many of the remaining typos. Of course, I am responsible for any surviving mistakes. I apologize to the reader in advance for any confusion that may arise due to them. And finally, I must confess that I have not always followed the suggestions made by the proofreaders. Some times due to lack of space and time, sometimes due to personal eccentricity. However, I hope that the value of this book has not been compromised as a result.

Part I

BACKGROUND

Chapter 1

Functions

1.1 Sets

In mathematics, the most primitive notion is that of a **set**. A set is a collection of objects with a common property, and these objects are called the **elements** of the set.

The set with no elements is called the **empty set** and is denoted by $\varnothing$ or $\{\}$.

Two sets A, B containing the same elements are called **equal**; we write $A = B$. If the sets are not equal, we write $A \neq B$ and we say that they are **different**.

If every element of the set A also belongs to the set B, we say that A is a **subset** of B, and we write $A \subseteq B$. If the set B has at least one element which does not belong to A, then we call A a **proper subset** of B, and we write $A \subset B$. Alternatively, one can use the notation $B \supseteq A$ and $B \supset A$; they are read "B is a **superset** of A" and "B is a **proper superset** of A" respectively. For example, the set of characters in this sentence is a proper subset of the set of characters on this entire page.

Let $A \subseteq S$. The **complement** of A relative to S is the set of elements $x \in S$ which do not belong to A. Usually this set is denoted by $A^{\complement}$ or $S \setminus A$.

Let $A, B \subseteq S$. The **union** of A and B, denoted by $A \cup B$, is defined as the set of $x \in S$ such that $x \in A$ or $x \in B$.

Let $A, B \subseteq S$. The **intersection** of A and B, denoted by $A \cap B$, is defined as the set of points $x \in S$ such that $x \in A$ and $x \in B$. Two sets A, B are called **disjoint** if $A \cap B = \varnothing$.

Let $A, B \subseteq S$. The **difference** $A - B$ is defined as the set of points $x \in A$ which do not belong to B.

The complement $S \setminus B$ can always be regarded as the difference of $S - B$. In this way we can use the universal notation $A - B$ to describe complements and differences.

We can take our discussion one level of abstraction higher by noting that given a set S we can form new sets whose *elements* are various subsets of S. Amongst these, a particularly interesting case is that of the set whose elements consist of *all* subsets of the set S. This set is known as the **power set** of S, and we denote this by $\mathcal{P}(S)$.

Finally, the **Cartesian product** of a collection of sets $\{A_1, A_2, ..., A_n\}$ is defined as the set whose elements consist of all n-tuples whose i-th entry is an element of the set A_i.

Symbolically, we write

$$A_1 \times A_2 \times \ldots \times A_n \;\equiv\; \{\,(a_1, a_2, \ldots, a_n),\; a_i \in A_i,\; i = 1, 2, \ldots, n\,\}\,. \tag{1.1}$$

If $A_1 = A_2 = \ldots = A_n \equiv A$, we usually write

$$\underbrace{A \times A \times \ldots \times A}_{n \text{ times}} \;=\; A^n\,.$$

Topological Concepts

Definition 1.1. Let $\mathcal{O}$ be a subset of the power set $\mathcal{P}(X)$ such that the following properties are true:
1. $\varnothing \in \mathcal{O}\,,\; X \in \mathcal{O}$.
2. The intersection of a finite number of sets in $\mathcal{O}$ belongs also to $\mathcal{O}$.
3. The union of any number of sets in $\mathcal{O}$ also belongs to $\mathcal{O}$.

The elements of $\mathcal{O}$ are then called the **open sets** of X with respect to the **topology** $\mathcal{T} = (X, \mathcal{O})$. The complement $O^{\complement} = X \setminus O$ of an open set O is defined as a **closed** set. The set X is called a **topological space** and its elements are called the **points** of X.

Definition 1.2. A set N is called a **neighborhood** of the point p if $p \in N$ and there exists an open set $O \in \mathcal{O}$ containing p which is contained in N, i.e. $O \subset N$.

In the above definition, a neighborhood may be open, closed, or neither. However, when we refer to a **neighborhood** of a point $a \in \mathbb{R}$, then we shall assume an open interval that contains a. A **symmetrical neighborhood** of a is a neighborhood centered at a, that is the interval $(a - \delta, a + \delta)$, where $\delta > 0$ is called the **radius** of the neighborhood. For our purposes it will suffice to consider symmetrical neighborhoods; any time we use the term 'neighborhood', we shall mean symmetrical neighborhood unless otherwise specified.

Given a space X, the following definitions provide some of the most interesting points and sets:

1. A point ℓ is called a **limit point** of the set S if every neighborhood $N(\ell)$ of ℓ contains at least one point of S different from ℓ. The set S' of all limit points of S is called the **derived set** of S.

2. A point i is called an **interior point** if there exists a neighborhood $N(i)$ of i that is contained in S. The set Int$S$ of all interior points of S is called the **interior set** of S.

3. A point e is called an **exterior point** of the set S if there exists a neighborhood $N(e)$ of e that is contained in $S^{\complement}$. The set Ext$S$ of all exterior points of S is called the **exterior set** of S.

4. A point b is called s **boundary point** of the set S if every neighborhood $N(b)$ of b is neither totally contained in S nor in $S^{\complement}$. The set ∂S of all boundary points of S is called the **boundary** of S.

5. A point i_0 is called an **isolated point** of the set S if it belongs to S and there is a neighborhood $N(i_0)$ of i_0 that does not contain any other point of S except i_0.

6. The smallest closed set $\overline{S}$ containing S is called the **closure** of S.

7. A subset S of X is called **dense** in X, if $\overline{S} = X$.

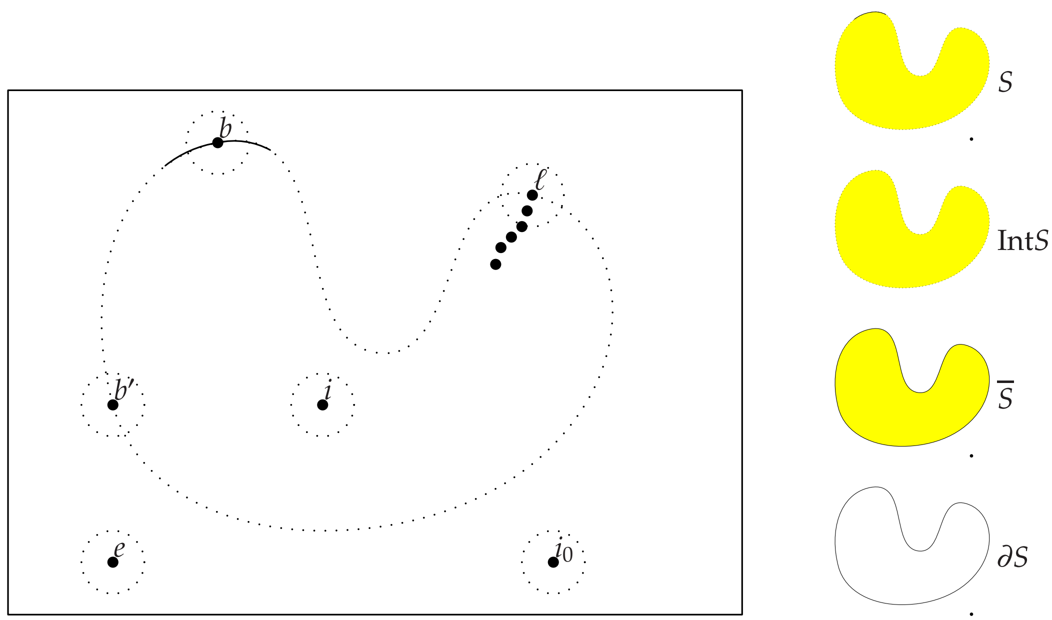

Figure 1.1: This figure gives a visualization of the points and sets we defined. The left side focuses on the terminology of points; the right side focuses on the terminology of sets. An interior point i of a set S is 'inside' S; an exterior point e is 'outside' (equivalently 'inside' the complement of the set); a boundary point lies exactly on the set that 'separates' the set from its complement and may or may not belong to the set such as b and b' respectively. An isolated point i_0 is a point of the set that happens to be 'far away' from any other point of the set; if it were not a point of S, it would qualify as an interior point of $S^{\complement}$. A limit point ℓ of S is the limit of a converging sequence of points in S.

There are many useful relations among the sets defined above. Let X be a topological space and A, B, O, C subsets of X. Then

1. The set C is closed iff[1] $\overline{C} = C$.

2. The set C is closed iff $C' \subseteq C$.

3. The set C is closed iff $\partial C \subseteq C$.

4. The set O is open iff $\text{Int}O = O$.

[1]The word *iff* stands for the phrase *if and only if*.

5. The set O is open iff $\partial O \cap O = \varnothing$.
6. $\overline{A} = A \cup A'$
7. $\overline{A \cup B} = \overline{A} \cup \overline{B}$
8. $\overline{\overline{A}} = \overline{A}$
9. $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$
10. $\overline{A - B} \subseteq \overline{A} - \overline{B}$
11. $\overline{X \setminus A} = \overline{X} - \mathrm{Int}A$
12. $\overline{A} = X \setminus \mathrm{Ext}A$
13. $\partial A = X \setminus (\mathrm{Int}A \cup \mathrm{Ext}A)$
14. $\partial A = \overline{A} \cap \overline{(X \setminus A)}$
15. $\partial A = \overline{A} - \mathrm{Int}A$

Example 1.1. An interesting set is the set of rational numbers $\mathbb{Q} = \{m/n\ ;\ m \in \mathbb{Z},\ n \in \mathbb{Z}^*\}$ as a subset of $\mathbb{R}$. Then $\overline{\mathbb{Q}} = \mathbb{R}$, $\partial\mathbb{Q} = \mathbb{R}$, $\mathrm{Int}\mathbb{Q} = \varnothing$. Since $\mathrm{Int}\mathbb{Q} \neq \mathbb{Q}$, $\mathbb{Q}$ is not open; since $\partial\mathbb{Q} \supset \mathbb{Q}$, $\mathbb{Q}$ is not closed either.

1.2 Relations

A **relation** R on the set $A \neq \varnothing$ is a subset of the Cartesian product $A \times A$, i.e. $R \subseteq A \times A$. Usually, instead of writing $(x, y) \in R$, we write xRy and we read '*x is related to y*'.

Definition 1.3. A relation $\sim$ on the set A is called an **equivalence relation** if it satisfies the following axioms for all $a, b, c \in A$:
1. *Reflexivity*: $a \sim a$,
2. *Symmetry*: $a \sim b$ implies $b \sim a$,
3. *Transitivity*: if $a \sim b$ and $b \sim c$ then $a \sim c$.

The expression $x \sim y$ is then read 'x is equivalent to y'.

Definition 1.4. Given an equivalence relation on A, the **equivalence class** of x is the set

$$[x] \ \equiv\ \{\, a \in A\ ;\ a \sim x \,\}\,.$$

Exercise 1.1. *Let $[x]$ and $[y]$ be two equivalence classes. Show that either $[x] = [y]$ or $[x] \cap [y] = \varnothing$.*

A **partition** of a set A is a family of disjoint subsets $\{A_i\}$ of A, such that $A = \cup_i A_i$. Using the last problem, we see that an equivalence relation in a set A defines a partition of A.

Example 1.2. In the set of integers $\mathbb{Z}$ we fix an integer k and we define the following equivalence relation:

$$\text{if } m, n \in \mathbb{Z} \quad \text{then} \quad (m \sim n \Leftrightarrow m - n \text{ is divisible by } k)\ .$$

Since any integer n can be uniquely written in the form

$$n \;=\; n' k + r\,, \quad 0 \leq r < k\,, \quad n', k, r \in \mathbb{Z}\,,$$

we conclude that two integers n, m belong to the same equivalence class only if they have the same residue r when divided by k. Therefore, this equivalence relation partitions the set of integers in k subsets

$$A_r \;=\; \{k\,n' + r \mid n' \in \mathbb{Z}\} \;=\; [r]\,, \qquad r = 0, 1, \ldots, k-1\ .$$

The set of equivalence classes is

$$\mathbb{Z}_k \;=\; \{[0], [1], \ldots [k-1]\}\ .$$

Often we write $\mathbb{Z}_k = \mathbb{Z}/\sim$ to show that $\mathbb{Z}_k$ is the set $\mathbb{Z}$ with the relation $\sim$ 'factored out'.

Definition 1.5. A relation $\preccurlyeq$ on the set A is called a **partial order relation** if it satisfies the following axioms:

1. Reflexivity: $\forall a,\ a \preccurlyeq a$.
2. Symmetry: for any a, b that are related by $a \preccurlyeq b$ and $b \preccurlyeq a$ then $a = b$,
3. Transitivity: for any a, b, c that are related by $a \preccurlyeq b$ and $b \preccurlyeq c$ then $a \preccurlyeq c$.

The expression $x \preccurlyeq y$ is then read 'x preceeds y'.

Notice that in an partial ordering of A, not all elements have to be related. However if we add the fourth axiom

4. Trichotomy: $\forall a, b \in A$, either $a \preccurlyeq b$ or $b \preccurlyeq a$,

then we have a **total order** .

Example 1.3. (a) Given a set A, let's construct $\mathcal{P}(A)$. Then, for $S, T \in \mathcal{P}(A)$ we define $S \preccurlyeq T$ if $S \subseteq T$. This relation defines a partial order in $\mathcal{P}(A)$ since sets that have no common elements cannot be ordered.

(b) The set of reals $\mathbb{R}$ equipped with $\leq$ is totally ordered.

Definition 1.6. A totally ordered set is said to be **well ordered** iff every non-empty subset of it has a least element.

Example 1.4. Every finite totally ordered set is well ordered. The set of integers, which has no least element, is an example of a set that is not well ordered.

1.3 Functions

The Notion of a Function

A **function** or **map** $f : A \to B$ between two sets A and B is a correspondence that assigns a *unique* element y of B to *every* element x of A. The set A is called the **domain** of f and the set B the **range** or **codomain** of f. For $x \in A$, the element y of B which corresponds to x is denoted by $f(x)$ and it is called the the **image** of x; we often write $y = f(x)$ or $x \overset{f}{\mapsto} y$. Also, x is called the **preimage** of y.

Comment. Although real-valued functions, $y \in \mathbb{R}$, of a real variable, $x \in \mathbb{R}$, will be the most common objects of our studies, the above definition and whatever follows is more general. The symbols x and y may be considered as vectors, matrices, or any other mathematical object. □

Example 1.5.

(a) If $A \neq \varnothing$, the function $\mathrm{id}_A : A \to A$ which maps each element to itself $\mathrm{id}_A(x) = x$ is called the **identity function**.

(b) If $A, B \neq \varnothing$, the function $f : A \to B$ which maps each element $x \in A$ to the same element $c \in B$, i.e. $f(x) = c,\ \forall x \in A$, is called a **constant function**. □

The set $\Gamma_f = \{(x, f(x))\ ,\ x \in A\} \subseteq A \times B$ is called the **graph** of f. We like to sketch this set to have a visual picture of a function (such as in Problem 1.1). However, for many functions such a pictorial representation may not be possible and we must treat the graph as a set.

Example 1.6. A function for which it is not easy to draw a sketch of the graph is the following:

$$f(x) = \begin{cases} 1\ , & x \in \mathbb{Q}\ , \\ 0\ , & x \in \mathbb{R} \smallsetminus \mathbb{Q}\ . \end{cases}$$ □

This is known as the **Dirichlet function**.

Problem 1.1. *Draw the graph of the function $f : \mathbb{R} \to \mathbb{R}$ if it satisfies the following two conditions:*
(a) $f(x+1) = f(x) - 2,\ \ \forall x \in \mathbb{R}$ *, and*
(b) $f(x) = x^2$, $x \in [0, 1)$.

Solution. Using induction it is easy to see that

$$f(x+n) = f(x) - 2n\ , \quad \forall n \in \mathbb{N}\ .$$

Let $y \in \mathbb{R}$ be an arbitrary real number. We can always write as $y = x + n$, where n is the integer part $\lfloor y \rfloor$ of y and x is the fractional part $\{y\}$ of y. Then, the previous equation gives

$$f(y) = f(\{y\}) - 2n\ , \quad n \leq y < n + 1\ .$$

Since the fractional part satisfies the inequality $0 \le \{y\} < 1$, $f(\{y\}) = \{y\}^2$ from (b). However, $\{y\} = y - n$ and we finally have

$$f(y) \;=\; (y-n)^2 - 2n\,, \quad n \le y < n+1\,.$$

The graph of this function is drawn in Figure 1.2. □

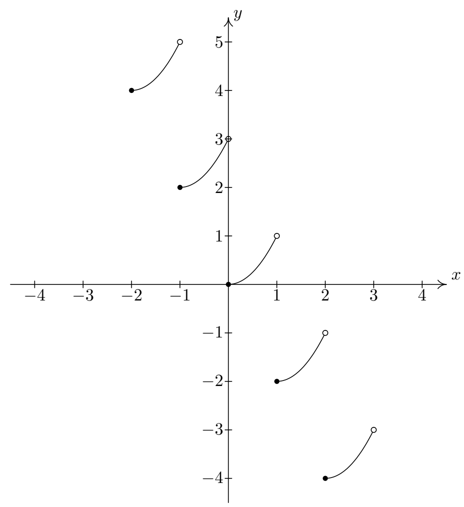

Figure 1.2: The graph of the function $f(x) = (x-n)^2 - 2n$, $x \in [n, n+1)$, $n \in \mathbb{Z}$.

Properties of Functions

Two functions $f : A \to B$ and $g : C \to D$ are said to be **equal** if $A = C$, $B = D$ and $f(x) = g(x)$ for all $x \in A$.

Let $f : A \to B$ be a function and $A' \subseteq A$. By $f(A')$ we denote the set of all images of the elements of A':

$$f(A') \;=\; \{y \in B \mid \exists x \in A',\; y = f(x)\}\,.$$

In particular, the set $f(A)$ is called the **image** of the function f; it is also denoted by $\mathrm{Im} f$. A function $f : A \to B$ is called **surjective** or **onto** if every element in the set B has a preimage under f in A. In other words, if $f(A) = B$. Obviously, every function can be surjective if we conveniently choose the set B to be the range of f.

Exercise 1.2. *Let $f : A \to B$ and A', B' be subsets of A. Prove that*

$$\begin{aligned} f(A' \cap B') &\subseteq f(A') \cap f(B')\,, \\ f(A' \cup B') &= f(A') \cup f(B')\,. \end{aligned}$$

□

A function $f : A \to B$ is called **injective** or **one-to-one** (1-1 for short) if distinct elements in A are mapped to distinct elements in B, i.e. if

$$\forall x_1, x_2 \in A,\ x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)\ .$$

Equivalently, f is injective if

$$\forall x_1, x_2 \in A,\ f(x_1) = f(x_2) \Rightarrow x_1 = x_2\ .$$

The function f is called **bijective** if it is both injective and surjective. Sometimes, we shall use the notation $f : A \rightarrowtail B$ to indicate an injective function, the notation $f : A \twoheadrightarrow B$ to indicate a surjective function, and the notation $f : A \rightarrowtail\!\!\!\!\twoheadrightarrow B$ to indicate a bijective function.

The **restriction** of a function $f : A \to B$ to a set $A' \subset A$ is the function $g : A' \to B$ where $g(x) = f(x),\ \forall x \in A'$. It is often denoted by $f\big|A'$. Equivalently, one can say that f is the **extension** of g to a set $A \supset A'$.

We say that the function[2] $f : A \to B$ is **strictly increasing** (**strictly decreasing**) if for any $x < y$ in A it is true that $f(x) < f(y)$ ($f(x) > f(y)$ respectively). We say that f is **increasing** or **non-decreasing** (**dereasing** or **non-increasing**) if for any $x < y$ in A it is true that $f(x) \leq f(y)$ ($f(x) \geq f(y)$ respectively). A strictly increasing or decreasing function is referred to as **strictly monotonic** while an increasing or decreasing function is referred to as **monotonic**.

We say that the function $f : A \to B$ is **upper bounded** (**lower bounded**) if there exists an $M \in B$ ($m \in B$) such that $f(x) < M$ ($m < f(x)$) for all $x \in A$. It is called **bounded** if it is both lower and upper bounded. For a lower bounded function, we call the greatest of its lower bounds the **infimum**; for an upper bounded function, we call the least of its bounds the **supremum**.

Operations on Functions

Given two functions $f : A \to B$ and $g : A \to C$ with the same domain, we define their sum $f + g$, difference $f - g$, product $f \cdot g$, and so on to be a function h defined on the same domain whose image $f(x)$ of $x \in A$ is the sum, difference, product, and so on of the images $f(x)$, $g(x)$ of the original functions.

Let $f : A \to B$ and $g : B \to C$ be two functions. We can define a new function by using f and g sequentially which maps $A \to C$. In particular, we define a new function $h : A \to C$, called the **composite** of f and g, by $h(x) = g(f(x)),\ \forall x \in A$. Often, the function h is written as $g \circ f$. The composition of functions satisfies associativity, that is $(f \circ g) \circ h = f \circ (g \circ h)$, where the domains are assumed such that all operations are well defined.

Comment. If $F(u, v)$ is a function of two real variables u, v with value $u+v$, $u-v$, uv, and so on, we notice that the composite function $F(f(x), g(x))$ is exactly the sum, difference, product and so on of the functions f and g. In other words, the functions defined by arithmetic operations between f and g are special cases of the composition of functions operation.

[2] The definitions that follow assume that A, B are sets with total order.

In the following the notation $f^n(x)$ will indicate[3] the composition of a function with itself: $f^2 = f \circ f$, $f^n = f \circ f^{n-1}$, $n > 2$. The quantity $f^n(x)$ is known as the n-th **iterate** of $f(x)$. Some care is required not to confuse this notation with the powers of $f(x)$: $(f)^2 = f \cdot f$, $(f)^n = f \cdot (f)^{n-1}$, $n > 2$. Usually I write $f(x)^2$ instead of $(f(x))^2$ to indicate powers of $f(x)$.

Inverse Function

A bijective function f provides unique association between the elements of the domain and range. This allows us to define another, closely related function which we shall temporarily call g with $g : B \to A$. Specifically, the bijectivity of f implies that each and every element $y \in B$ is associated to an element $x \in A$ by demanding $f(x) = y$. The function g carries out this association, i.e. $g : B \to A\ ;\ y \mapsto x$ by the rule $x = g(y) \Leftrightarrow y = f(x)$. Notice that this function is just the function f in reverse, and hence is called the **inverse function** of f and is typically denoted by f^{-1} instead of g. Obviously f^{-1} is also bijective.

Theorem 1.1. *For a bijective function f there is a unique inverse function f^{-1}.*

Comparing the graphs of f and f^{-1},

$$\begin{aligned} \Gamma_f &= \{(x, y),\ y = f(x),\ x \in A\}\,, \\ \Gamma_{f^{-1}} &= \{(y, x),\ x = f^{-1}(y),\ y \in B\} \end{aligned}$$

we notice that they are symmetric under reflection on the line $y = x$.

Exercise 1.3. *Let $f : A \rightarrowtail\!\!\!\!\twoheadrightarrow B$ and $g : B \rightarrowtail\!\!\!\!\twoheadrightarrow C$ be two bijective functions. Prove the following properties:*
(a) $(f^{-1})^{-1} = f$,
(b) $f^{-1} \circ f = \mathrm{id}_A$,
(c) $f \circ f^{-1} = \mathrm{id}_B$,
(d) $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. □

Problem 1.2 (IMO 1973). *Let G be the set of functions $f : \mathbb{R} \to \mathbb{R}$ of the form $f(x) = a\,x + b$, where a and b are real numbers and $a \neq 0$. Suppose that G satisfies the following conditions:*
(a) If $f, g \in G$ then $g \circ f \in G$.
(b) If $f \in G$ then $f^{-1} \in G$.
(c) For each $f \in G$ there exists a number $x_f \in \mathbb{R}$ such that $f(x_f) = x_f$.
Prove that there exists a number $k \in \mathbb{R}$ such that $f(k) = k$ for all $f \in G$.

Solution. If $a = 1$ for a function $f(x) = x + b$, then condition (c) requires that $b = 0$. In this case, all points k of $\mathbb{R}$ satisfy $f(k) = k$. Therefore we need to show it for $a \neq 1$.

Let $a, a' \neq 1$ and

$$f(x) = a\,x + b\,, \quad g(x) = a'\,x + b'\,,$$

[3] I break this rule only for the trigonometric functions: $\sin^2 x$ stands for $(\sin x)^2$, not $\sin(\sin x)$. This tradition is so strong and widespread that I found impossible to reverse it.

be two functions in G. Then, condition (c) requires that there are two points x_f and x_g (not necessarily distinct) such that

$$f(x_f) = x_f \Rightarrow x_f = \frac{b}{a-1},$$

and

$$g(x_g) = x_g \Rightarrow x_g = \frac{b'}{a'-1}.$$

According to condition (b), both

$$f^{-1} = \frac{1}{a}x - \frac{b}{a}, \quad g^{-1} = \frac{1}{a'}x - \frac{b'}{a'}$$

are in G. Then, according to condition (a),

$$f \circ g(x) = aa'\,x + ab' + b,$$

and

$$f^{-1} \circ g^{-1}(x) = \frac{1}{aa'}x - \frac{b' + ba'}{aa'},$$

and

$$f \circ g \circ f^{-1} \circ g^{-1} = x + (ab' + b) - (b' + ba')$$

are also elements of G. Since there is a x_0 for this function such that $f \circ g \circ f^{-1} \circ g^{-1}(x_0) = x_0$, we conclude that

$$(ab' + b) - (b' + ba') = 0 \Rightarrow \frac{b}{1-a} = \frac{b'}{1-a'} \Rightarrow x_f = x_g. \qquad \square$$

If $B' \subseteq B$, by $f^{-1}(B')$ we denote the set of all preimages of the elements of B':

$$f^{-1}(B') = \{x \in A \mid \exists y \in B',\ x = f^{-1}(y)\}.$$

Exercise 1.4. *Let A' be a subset of A and B', B'' be two subsets of B. Prove that*

$$\begin{aligned} f^{-1}(f(A')) &\supseteq A', \\ f(f^{-1}(B')) &\subseteq B', \\ f^{-1}(B' \cap B'') &= f^{-1}(B') \cap f^{-1}(B''), \\ f^{-1}(B' \cup B'') &= f^{-1}(B') \cup f^{-1}(B''). \end{aligned} \qquad \square$$

1.4 Limits & Continuity

Limits

A function $f : \mathbb{R} \to \mathbb{R}$ is said to have a **limit** $L \in \mathbb{R}$ at $a \in \mathbb{R}$, denoted by $\lim\limits_{x\to a} f(x) = L$ or $f(x) \underset{x\to a}{\longrightarrow} L$, if for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - L| < \varepsilon,\ \forall x \in \mathbb{R} \text{ and } 0 < |x - a| < \delta\ .$$

Using neighborhoods, the limit of f at a point a may be restated as follows: A function $f : \mathbb{R} \to \mathbb{R}$ is said to have a **limit** $L \in \mathbb{R}$ at $a \in \mathbb{R}$ if, given any neighborhood J of the point L, there exists a neighborhood I of the point a such that $f(I) \subset J$.

The definition of the limit can easily be modified to consider **one-sided limits**, that is limits of $f(x)$ as x approaches a from the right (we write $\lim\limits_{x\to a^+} f(x)$ and we call it the **right-hand limit** of $f(x)$) or as x approaches a from the left (we write $\lim\limits_{x\to a^-} f(x)$ and we call it the **left-handed limit** of $f(x)$).

Exercise 1.5. *If* $\lim\limits_{x\to a} f(x) = \ell_1$ *and* $\lim\limits_{x\to a} g(x) = \ell_2$ *then prove the following properties:*

(a) $\lim\limits_{x\to a}(\lambda f(x)) = \lambda\, \ell_1$, *for any* $\lambda \in \mathbb{R}$.

(b) $\lim\limits_{x\to a}(f(x) + g(x)) = \ell_1 + \ell_2$.

(c) $\lim\limits_{x\to a} f(x)\, g(x) = \ell_1 \ell_2$.

(d) $\lim\limits_{x\to a} \frac{f(x)}{g(x)} = \frac{\ell_1}{\ell_2}$ *provided that* $g(x) \neq 0$ *and* $\ell_2 \neq 0$. □

The following two theorem give two additional important properties of the limit.

Theorem 1.2. *If* $f(x) \le g(x)$ *in some neighborhood of* a *and the limits of* f, g *as* $x \to a$ *exist, then*

$$\lim_{x\to a} f(x) \le \lim_{x\to a} g(x)\ .$$

Theorem 1.3 (Squeeze Theorem)**.** *If* $f(x) \le g(x) \le h(x)$ *in some neighborhood of* a *and*

$$\lim_{x\to a} f(x)\ =\ \lim_{x\to a} h(x)\ =\ \ell\ ,$$

then

$$\lim_{x\to a} g(x)\ =\ \ell\ .$$

A sequence $\{x_n\}_{n\in\mathbb{N}}$ of real numbers is said to **converge** to a limit $L \in \mathbb{R}$, denoted by $\lim\limits_{n\to\infty} x_n = L$ or $a_n \to L$, if for any $\varepsilon > 0$, there exists n_0 such that $|x_n - L| < \varepsilon$ for all $n \ge n_0$. Similar results with those of functions apply to sequences.

Continuity

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be **continuous** at $a \in \mathbb{R}$ if $\lim_{x\to a} f(x) = f(a)$. Moreover, f is called continuous on an open interval $I \subseteq \mathbb{R}$ if f is continuous at every $x \in I$. Using right-hand and left-hand limits, we can extend the concept of continuity to *continuity from right* or *continuity from left* respectively.

Also, using the open sets of a topology, we can extend the concept of continuity to any function between two topological spaces: The function $f : X \to Y$ is said to be **continuous** at $a \in X$ if, given any open set J of the point $f(a)$, there exists an open set I of the point a such that $f(I) \subset J$. The function f is continuous on X if it is continuous at every point $x \in X$.

Theorem 1.4. *A function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a iff $\lim_{n\to\infty} f(x_n) = f(a)$ for every sequence $\{x_n\}$ converging to a.*

A function that is not continuous at the point a is called **discontinuous** at a. Such a discontinuity can occur for one of the following reasons:

(a) The right-hand and left-hand limits are equal but different from the value of the function.

(b) The right-hand and left-hand limits exist but they have different values.

(c) The right-hand or left-hand limit does not exist.

In the first case, we can re-define the function at a and the function becomes continuous. The discontinuity is thus called **removable discontinuity** In the other two cases, the discontinuity cannot be remove and thus are called **irremovable discontinuities** of **first kind** (or **jump discontinuity**) and **second kind** (or **essential discontinuity**) respectively.

The function of Problem 1.1 has jump discontinuities at all integer points. The function

$$f(x) = \begin{cases} \frac{1}{x^2}\,, & x \neq 0\,, \\ 0\,, & x = 0\,, \end{cases}$$

has an essential discontinuity at $x = 0$. The function

$$f(x) = \sin\frac{1}{x}\,, \quad x \neq 0$$

oscillates violently from -1 to 1 close to $x = 0$ and it cannot have a limit as x approaches this point — an essential discontinuity at $x = 0$. Introductory calculus texts call the discontinuity at $x = 0$ of this function an **oscillating discontinuity**.

Finally for the function

$$f(x) = x\sin\frac{1}{x}\,, \quad x \neq 0\,,$$

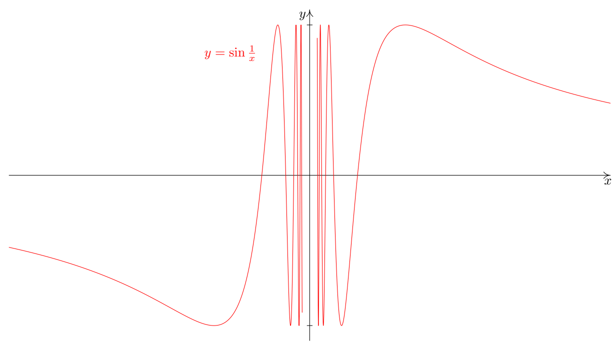

Figure 1.3: The graph of the function $f(x) = \sin\frac{1}{x}$. The function is continuous at all points $x = 0$ but at $x = 0$ its has an essential discontinuity. To see this, take the sequences $a_n = \frac{2}{(4n+1)\pi}$ and $b_n = \frac{2}{(4n-1)\pi}$, both converging to 0. However $\sin a_n = \sin\frac{\pi}{2} = 1$ and $\sin b_n = \sin\left(-\frac{\pi}{2}\right) = -1$, thus violating Theorem 1.4.

the limit as $x \to 0$ exists and it is zero. Since $x = 0$ does not belong to the domain of $f(x)$, as it stands, $f(x)$ is discontinuous at $x = 0$. However, if we define

$$f(x) = \begin{cases} x\sin\frac{1}{x}\,, & x \neq 0\,, \\ 0\,, & x = 0\,, \end{cases}$$

then $f(x)$ becomes continuous at all points. The initial discontinuity at $x = 0$ is a removable discontinuity.

The previous discussion may have created the impression that functions are 'mostly' continuous and they may have only a 'small' set of points where discontinuities appear. However, such an impression is not correct. The Dirichlet function of Example 1.6 is discontinuous *everywhere* and the function

$$f(x) = \begin{cases} x\,, & x \in \mathbb{Q}\,, \\ 0\,, & x \in \mathbb{R} \setminus \mathbb{Q}\,, \end{cases}$$

which is perhaps the simplest modification of the Dirichlet function, is continuous only at one point — at $x = 0$.

Exercise 1.6. *Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be two continuous functions at $x = a$. Prove the following properties:*

(a) $f + g$ is continuous at $x = a$.
(b) fg is continuous at $x = a$.
(c) f/g is continuous at $x = a$ provided that $g(a) \neq 0$.
(d) $f \circ g$ is continuous at $x = a$. □

A very useful theorem for continuous functions with many applications is the following.

Theorem 1.5 (Bolzano)**.** *If the function $f(x)$ is continuous on the interval $[a, b]$ and $f(a)f(b) < 0$, then there exists a point $\xi \in (a, b)$ such that $f(\xi) = 0$.*

That is, the graph of continuous function cannot change sign unless it crosses the x-axis. Bolzano's theorem can be extended to

Theorem 1.6 (Intermediate Value Theorem)**.** *If the function $f(x)$ is continuous on the interval $[a, b]$ then it takes on every value between $f(a)$ and $f(b)$.*

That is, a horizontal line crossing the y-axis at a point between $f(a)$ and $f(b)$, it necessarily crosses the graph of a continuous function $f(x)$ with domain $[a, b]$.

If a function is discontinuous, then the previous two theorems do not hold. For example, the function, in Bolzano's theorem, may jump from positive to negative values (or the other way around) without crossing the horizontal axis such as the function of Problem 1.1.

1.5 Differentiation

First Derivative

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be **differentiable** at a point a if the

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a} \;=\; \lim_{h\to 0} \frac{f(h + a) - f(a)}{h}$$

exists. This limit is then called the **derivative** of f at $x = a$ and denoted by $f'(a)$ or $df(x)/dx|_a$. Moreover, f is called differentiable on an open interval $I \subseteq \mathbb{R}$, if f is differentiable at every $x \in I$. Using right-hand and left-hand limits, we can extend the concept of differentiation to *differentiation from right* or *differentiation from left* respectively. Geometrically, the derivative $f'(a)$ gives the slope of the tangent line at the point $(a, f(a))$ of Γ_f.

Theorem 1.7. *If a function $f(x)$ is differentiable at the point a, then it is continuous at the point a.*

Differentiability of the function $f(x)$ at x is equivalent to the continuity of the function

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

at $h = 0$, with the variable x playing the role of a parameter specifying the point of interest. Therefore any kind of discontinuity at $h = 0$, except removable discontinuity, signals problems. For the function $f(x) = |x|$,

$$F(h) \;=\; \frac{|x + h| - |x|}{h} \,.$$

For $x = 0$,

$$F(h) \;=\; \frac{|h|}{h} \;=\; \mathrm{sgn}(h)\,, \quad h \neq 0\,.$$

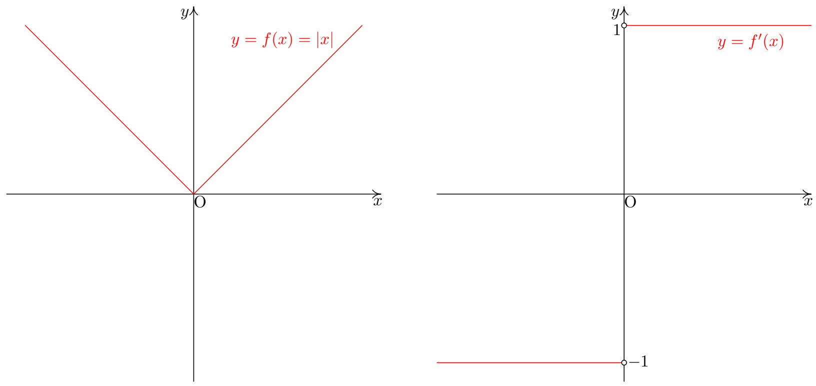

Figure 1.4: Left: The graph of the absolute value function has a 'spike' at $x = 0$. Right: The derivative of the absolute value function is the sign function for $x \neq 0$. At $x = 0$, the left and right derivatives exhibit a jump discontinuity and thus the absolute value function has no derivative at $x = 0$.

This is a discontinuous function with a finite jump discontinuity at $x = 0$. Therefore $f(x) = |x|$ does not have a derivative at $x = 0$. The jump discontinuity in $F(h)$ appears as a 'spike' in the graph of $f(x)$.

The function

$$f(x) = \begin{cases} x \sin \frac{1}{x}\,, & x \neq 0\,, \\ 0\,, & x = 0\,, \end{cases}$$

implies

$$F(h) = \sin \frac{1}{h}\,,$$

at $x = 0$ which has an essential discontinuity and therefore $f(x)$ has no derivative at this point. Finally the function

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x}\,, & x \neq 0\,, \\ 0\,, & x = 0\,, \end{cases}$$

implies a continuous function

$$F(h) = h \sin \frac{1}{h}\,,$$

at $x = 0$ and therefore $f'(0)$ exists and it is equal to $f'(0) = 0$.

Continuous functions may appear to fail to be differentiable only are only at a 'small' set of points. Such a point of view was held by mathematicians until Weierstrass[4] explicitly showed to be an incorrect one. We present the Weierstrass function after presenting some additional results on the first derivative.

[4]Bolzano seems to have done so before Weierstrass but his work was not paid attention to.

Exercise 1.7. *If f and g are differentiable at x and c is a constant, prove the following properties:*
(a) $(f+g)'(x) = f'(x) + g'(x)$.
(b) $(cf)'(x) = cf'(x)$.
(c) $(fg)'(x) = f(x)g'(x) + f'(x)g(x)$.
(d) $(f/g)'(x) = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$, *for* $g(x) \neq 0$.
(d) $(f(g(x))' = f'(g(x))\, g'(x)$. □

The following three theorems (each generalizing the preceding one) are among the most useful theorems for differentiable functions.

Theorem 1.8 (Rolle)**.** *If $f(x)$ is continuous on $[a,b]$, differentiable on (a,b) and $f(a) = f(b) = 0$, then there is a $\xi \in (a,b)$ such that $f'(\xi) = 0$.*

That is, if f is continuous, then Γ_f has a local extremum between any two roots. (See the section on the application of derivatives.)

Theorem 1.9 (Lagrange's Mean Value Theorem)**.** *If $f(x)$ is continuous on $[a,b]$ and differentiable on (a,b), then there is a $\xi \in (a,b)$ such that*

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

That is, if f is continuous, then for every secant line of Γ_f, there is at least one parallel line, tangent to Γ_f.

Theorem 1.10 (Cauchy's Mean Value Theorem)**.** *If $f(x), g(x)$ are continuous on $[a,b]$, differentiable on (a,b), and $g(x) \neq 0$ for any $x \in (a,b)$, then there is a $\xi \in (a,b)$ such that*

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b)-f(a)}{g(b)-g(a)} .$$

That is, if f, g are continuous, then there is always a point ξ for which the ratio of the slopes of the tangent lines to Γ_f and Γ_g is equal to the ratio of the slope of the secant lines crossing the graphs at $(a, f(a))$, $(b, f(b))$ and $(a, g(a))$, $(b, g(b))$ respectively.

Example 1.7 (Weierstrass function)**.** Let

$$f(x) = \sum_{k=0}^{\infty} a^k \cos(b^k \pi x) ,$$

where $0 < a < 1$ and b is a positive odd integer. I will leave the proof of continuity to the reader (or just simply accept it) and I will examine the differentiability of this function. Also, since the initial function is defined as an infinite sum, he may accept as valid any required re-arragement in the terms.

For the function $F(h)$, let's write $F = A_n + R_n$ where A_n is the sum of the first n terms

$$A_n \;=\; \sum_{k=0}^{n-1} a^k \,\frac{\cos[b^k\pi(x+h)] - \cos(b^k\pi x)}{h}\;,$$

and R_n is the remainder

$$R_n \;=\; \sum_{k=n}^{\infty} a^k \,\frac{\cos[b^k\pi(x+h)] - \cos(b^k\pi x)}{h}\;.$$

Applying Lagrange's mean value theore for the function $\cos(b^k\pi x)$ on $[x, x+h]$, we see that there is a $\xi \in (x, x+h)$ such that

$$b^k\pi\,\sin(b^k\pi\xi) \;=\; \frac{\cos[b^k\pi(x+h)] - \cos(b^k\pi x)}{h}\;.$$

Then, by the triangle inequality, and the above,

$$\begin{aligned}
|A_n| \;&\le\; \sum_{k=0}^{n-1} a^k \left|\frac{\cos[b^k\pi(x+h)] - \cos(b^k\pi x)}{h}\right| \\
&=\; \pi \sum_{k=0}^{n-1} (ab)^k \left|\sin(b^k\pi\xi)\right| \\
&\le\; \pi \sum_{k=0}^{n-1} (ab)^k \;=\; \pi\,\frac{(ab)^n - 1}{ab-1} < \pi\,\frac{(ab)^n}{ab-1}\;,
\end{aligned}$$

assuming that $ab > 1$.

Now we set $b^n x = N_n + f_n$, where N_n is an 'integer' and f_n a 'fractional' part, the latter defined here in the interval $[-1/2, 1/2]$. If $h = (1 - f_n)/b^n$, then $2b^n/3 \le 1/h \le 2b^n$. Also, for any $k \ge n$

$$\begin{aligned}
\cos[b^k\pi(x+h)] \;&=\; (-1)^{N_n+1}\;, \\
\cos(b^k\pi x) \;&=\; (-1)^{N_n+1}\,\cos(b^{k-n}\pi f_n)\;.
\end{aligned}$$

Therefore, for any $k \ge n$

$$\begin{aligned}
|R_n| \;&=\; \left|\sum_{k=n}^{\infty} a^k\,\frac{1 - \cos(b^{k-n}\pi f_n)}{h}\right| \\
&>\; \left|a^n\,\frac{1-\cos(\pi f_n)}{h}\right| \ge \frac{2(ab)^n}{3}\,[1 - \cos(\pi f_n)] > \frac{2(ab)^n}{3}\;.
\end{aligned}$$

Since $|F| = |A_n + R_n|$, we have

$$|F| \ge |R_n| - |A_n| > (ab)^n\,\frac{2ab - (2+3\pi)}{3(ab-1)}\;.$$

And this is true for any n. If $2ab > 2 + 3\pi$, the right hand side grows without bound as $n \to \infty$. Therefore, F cannot be continuous implying that f cannot be differentiable. Since x has been left unspecified, the result is for any x in the domain of f. □

Higher Derivatives

The derivative function $f'(x)$ of a function $f(x)$ may or may not be differentiable. If $f''(x)$, $f'''(x), \ldots, f^{(n)}(x)$ exist, then the function $f(x)$ is called **n-times differentiable**.

Theorem 1.11 (Leibniz)**.** *If f, g are n-times differentiable, then*

$$\frac{d^n}{dx^n} f(x)g(x) \;=\; \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x)\, g^{(n-k)}(x)\,.$$

Theorem 1.12 (Taylor)**.** *If f is $(n+1)$-times differentiable with continuous derivatives on $[a, x]$, then there exists some $\xi \in (a, x)$ such that*

$$f(x) \;=\; \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(x-a)^k + R_n\,,$$

with

$$R_n \;=\; \frac{f^{(n+1)}(\xi)}{(n+1)!}\,(x-a)^{n+1}\,.$$

The quantity R_n is known as the **remainder of order** n. If the derivative $f^{(n)}$ exists for any n, then the function is said to be **smooth** or simply infinite-times differentiable. For a smooth function, one can construct the so-called **Taylor series**

$$\sum_{n=1}^{\infty} \frac{1}{n!}\, f^{(n)}(a)\,(x-a)^n$$

which does not necessarily converge. If it converges to $f(x)$ on some neighborhood I of a with radius r, then we say that the function $f(x)$ is **analytic** at the point a with radius of convergence r. If $f(x)$ is analytic for all points of I, we say that f analytic on I.

Applications of Derivatives

Definition 1.7. (a) We say that the point x_0 is an **absolute maximum** of the function $f\big|D$ if $f(x) \leq f(x_0)$, for all $x \in D$. We say that the point x_0 is a **local maximum** of the function $f\big|D$ if $f(x) \leq f(x_0)$, for all x in some neighborhood of x_0.

(b) We say that the point x_0 is an **absolute minimum** of the function $f\big|D$ if $f(x) \geq f(x_0)$, for all $x \in D$. We say that the point x_0 is a **local minimum** of the function $f\big|D$ if $f(x) \geq f(x_0)$, for all x in some neighborhood of x_0.

(c) Absolute (local) minima and absolute (local) maxima are called collectively absolute (local) **extrema**.

Theorem 1.13 (Extreme Value Theorem)**.** *If f is continuous on $[a, b]$, then f attains an absolute maximum and absolute minimum at some points of $[a, b]$.*

Problem 1.3 (Bulgaria 2000)**.** *Let*

$$f(x) = \frac{x^2 + 4x + 3}{x^2 + 7x + 14}, \quad g(x) = \frac{x^2 - 5x + 10}{x^2 + 5x + 20}.$$

(a) Find the maximum value of $f(x)$.
(b) Find the maximum value of $g(x)^{f(x)}$.

Solution. The quadratic polynomials $x^2 + 7x + 14$, $x^2 - 5x + 10$, $x^2 + 5x + 20$ have a negative discriminant and therefore they are positive for all values of x. Among other things, this implies that $g(x) > 0$ and therefore the function $g(x)^{f(x)}$ is well defined.

The quadratic polynomial $x^2 + 4x + 3$ has roots -1 and -3. Therefore $f(x)$ is positive for $x \in (-\infty, -3) \cup (-1, +\infty)$ and negative for $x \in (-3, -1)$.

Finally,

$$g(x) = \frac{x^2 - 5x + 10}{x^2 + 5x + 20} = \frac{(x^2 + 5x + 20) - 10(x + 1)}{x^2 + 5x + 20} = 1 - 10\,\frac{x + 1}{x^2 + 5x + 20}.$$

That is, $g(x) \leq 1$ for $x \geq -1$ and $g(x) \geq 1$ for $x \leq -1$.

(a) The maximum value of $f(x)$ is 2. Indeed,

$$\begin{aligned} f(x) \leq 2 \quad &\Leftrightarrow \quad \frac{x^2 + 4x + 3}{x^2 + 7x + 14} \leq 2 \\ &\Leftrightarrow \quad x^2 + 4x + 3 \leq 2x^2 + 14x + 28 \\ &\Leftrightarrow \quad 0 \leq x^2 + 10x + 25 = (x + 5)^2. \end{aligned}$$

The maximum value is attained for $x = -5$.

(b) As in part (a),

$$g(x) \leq 3 \;\Leftrightarrow\; 0 \leq (x + 5)^2.$$

The function $g(x)$ attains a maximum value of 3 for $x = -5$.

Since $g(x) > 0$, $\ln g(x) > \ln 3$. For $x \in (-\infty, -3]$,

$$f(x)\, \ln g(x) \leq 2 \ln 3 \;\Rightarrow\; g(x)^{f(x)} \leq 9.$$

For $x \in [-3, -1]$,

$$g(x) \geq 1 \;\Rightarrow\; g(x)^{|f(x)|} \geq 1 \;\Rightarrow\; g(x)^{f(x)} \leq 1.$$

For $x \in [-1, +\infty]$,

$$g(x) \leq 1 \;\Rightarrow\; g(x)^{f(x)} \leq 1.$$

So, the maximum value of $g(x)^{f(x)}$ is 9 attained at $x = -5$. □

The above problem demonstrates that quite some work is necessary to find extrema of a function by algebraic methods. In addition, having numbers that always conspire to simplify the calculations is impossible. The study of extrema is systematized by the derivatives of a function. We shall call **critical points** of a function f those points c at which $f'(c) = 0$ or f' does not exist.

Theorem 1.14. *Extreme values of a function occur at critical points and endpoints.*

However, a critical point may not necessarily be a point of extreme value. The next two theorems provide criteria to verify when a critical point leads to an extremum.

Theorem 1.15 (First Derivative Test)**.** *If c is a critical point of f, then f has a local extremum at c if the derivative f' changes sign as it crosses c. In particular, it is a local maximum if $f' > 0$ for $x < c$ and $f' < 0$ for $x > c$ and it is a local minimum if $f' < 0$ for $x < c$ and $f' > 0$ for $x > c$.*

Corollary. *(a) A function is strictly increasing (respectively increasing) if $f' > 0$ (respectively $f' \geq 0$).*

(b) A function is strictly decreasing (respectively decreasing) if $f' < 0$ (respectively $f' \leq 0$).

Theorem 1.16 (Second Derivative Test)**.** *If c is a critical point of f, then f has a local minimum at c if $f''(c) > 0$ or has a local maximum if $f''(c) < 0$.*

Exercise 1.8. *If $f''(x_0) = f'''(x_0) = \cdots = f^{(2n)}(x_0) = 0$, but $f^{(2n+1)}(x_0) \neq 0$, discuss the behavior of f in the neighborhood of x_0. The point x_0 is called a **point of inflection**.*

1.6 Solved Problems

In this section we present some additional solved problems on functions. Of course this subject is immense. We only choose to work out some sample problems related to the topic of the book to set the stage for the following chapters.

Problem 1.4 (Singapore 2002)**.** *Let $f(x)$ be a function which satisfies*

$$f(29 + x) = f(29 - x)\,, \quad \forall x \in \mathbb{R}\,.$$

If $f(x)$ has exactly three real roots α, β, γ, determine the value of $\alpha + \beta + \gamma$.

Solution. Consider the coordinate system $O'XY$ that is obtained from the Oxy by a change of variables: $Y = y$, $X = x - 29$. This is just a simple translation of the origin $O(0, 0)$ to $O'(29, 0)$. The points $x_\pm = 29 \pm x$, in the new system become $X_\pm = \pm x$, and the given functional equation becomes

$$f(x) = f(-x)\,.$$

Therefore, the graph of the function $f(x)$ is even with respect to $X = 0$ and, as such, the roots much be located symmetrically with respect this line too. Since there are three roots, one must be located at $X = 0$, that is $\alpha' = 0$. The other two must be at some $\beta' = x_0$ and $\gamma' = -x_0$. Returning to the original system, $\alpha = 29$, $\beta = x_0 + 29$, and $\gamma = -x_0 + 29$. Then $\alpha + \beta + \gamma = 87$. □

Problem 1.5 ([13], Problem 7). *Let* $f_0(x) = \frac{1}{1-x}$*, and* $f_n(x) = f_0(f_{n-1}(x))$*,* $n = 1, 2, 3, \ldots$ *. Evaluate* $f_{1976}(1976)$*.*

Solution. We notice that

$$\begin{aligned} f_0(x) &= \frac{1}{1-x}\,, \\ f_1(x) &= f_0(f_0(x)) = \frac{1-x}{-x}\,, \\ f_2(x) &= f_0(f_1(x)) = x\,, \\ f_3(x) &= f_0(f_2(x)) = \frac{1}{1-x} = f_0(x)\,. \end{aligned}$$

From these results we conclude that

$$f_{3k+r}(x) = f_r(x)\,, \quad k = 0, 1, 2, \ldots\,, \quad r = 0, 1, 2\,.$$

Therefore

$$f_{1976}(x) = f_2(x)\,,$$

and, in particular, $f_{1976}(1976) = f_2(1976) = 1976$. □

Problem 1.6. *Let* $f(x) = x^2 - 2$ *with* $x \in [-2, 2]$*. Show that the equation*

$$f^n(x) = x$$

has 2^n *real roots.*

Solution. Since $x \in [-2, 2]$ we set $x = 2\cos\theta$, $0 \le \theta \le 2\pi$. Then

$$\begin{aligned} f(\cos\theta) &= 2[2\cos^2\theta - 1] = 2\cos(2\theta)\,, \\ f(f(\cos\theta)) &= [2\cos(2\theta)]^2 - 2 = 2[2\cos^2(4\theta) - 1] = 2\cos(4\theta)\,, \\ &\ldots \end{aligned}$$

By induction, we easily verify that

$$f^n(\cos\theta) \;=\; 2\cos(2^n\theta)\;.$$

The given equation, in the new notation, becomes

$$2\cos(2^n\theta) \;=\; 2\cos\theta\;,$$

with solutions $2^n\theta = 2k\pi \pm \theta$, $k \in \mathbb{Z}$ or

$$\theta_k^- \;=\; k\,\frac{2\pi}{2^n-1}\;,\quad \theta_k^+ \;=\; k\,\frac{2\pi}{2^n+1}\;,\quad k \in \mathbb{Z}\;.$$

The distinct solutions are those for which $0 \le \theta < 2\pi$. Therefore,

$$\begin{aligned}\theta_k^- &= k\,\frac{2\pi}{2^n-1}\;,\quad k = 0,1,\ldots,2^{n-1}-1\;,\\ \theta_k^+ &= k\,\frac{2\pi}{2^n+1}\;,\quad k = 1,\ldots,2^{n-1}\;.\end{aligned}$$

Counting, these are exactly 2^n in number. □

This problem actually appeared as one of the problems of the IMO 1976:

Problem 1.7 (IMO 1976)**.** *Let $P_1(x) = x^2 - 2$ and $P_j(x) = P_1(P_{j-1}(x))$ for $j = 2, 3, \cdots$. Show that for any positive integer n, the roots of the equation $P_n(x) = x$ are real and distinct.*

In this statement the domain is not specified to be $[-2, 2]$. The problem is actually taken from the theory of *orthogonal polynomials*. In particular, the functions $T_N(\cos\theta) = \cos(N\theta)$ are known as the **Chebychev polynomials of the first kind**. When written in terms of the variable $x = \cos\theta$, they are indeed polynomials as one can easily verify. The function $f^n(x)$ of the problem is just $2T_{2^n}(x)$.

Here is a related problem that was proposed for the IMO 1978:

Problem 1.8 (IMO 1978 Longlist)**.** *Given the expression*

$$P_n(x) \;=\; \frac{1}{2^n}\left[\left(x + \sqrt{x^2-1}\right)^n + \left(x - \sqrt{x^2-1}\right)^n\right]\;,$$

prove that $P_n(x)$ satisfies the identity

$$P_n(x) - x\,P_{n-1}(x) + \frac{1}{4}\,P_{n-2}(x) \;=\; 0\;,$$

and that $P_n(x)$ is a polynomial in x of degree n.

The Chebychev polynomials $T_n(x)$ satisfy the recursion relation

$$T_n(x) - 2xT_{n-1}(x) + T_{n-2}(x) = 0 ,$$

and are given by the expression

$$T_n(x) = \frac{1}{2}\left[\left(x + i\sqrt{1-x^2}\right)^n + \left(x - i\sqrt{1-x^2}\right)^n\right] .$$

Obviously the proposed problem is a multiplicative rewriting of the Chebychev polynomials:

$$P_n(x) = 2^{n-1}\,T_n(x) .$$

You can try to solve it anyway without this information.

The IMO 1976 problem, in the way stated, is not immediately related to the Chebychev polynomials, and it takes some time — even for the more experienced people — to make the connection[5]. However, the IMO 1978 problem is a straightforward routine exercise from the college textbooks. In any case, now you know the solution to the following problem given in the Swedish mathematical olympiad of 1996:

Problem 1.9 (Sweden 1996)**.** *For all integers $n \geq 1$ the functions p_n are defined for $x \geq 1$ by*

$$p_n(x) = \frac{1}{2}\left[\left(x + \sqrt{x^2-1}\right)^n + \left(x - \sqrt{x^2-1}\right)^n\right] .$$

Show that $p_n(x) \geq 1$ and that $p_{mn}(x) = p_m(p_n(x))$.

Again, you can try to solve it anyway, independently of what we have presented here.

The following problem and its solution shares several common ideas with the solution of the previous problem.

Problem 1.10 (Turkey 1998)**.** *Let $\{a_n\}$ be the sequence of real numbers defined by $a_1 = t$ and*

$$a_{n+1} = 4a_n\,(1 - a_n) , \quad n \geq 1 .$$

For how many distinct values of t do we have $a_{1998} = 0$?

Solution. We define the function $f(x) = 4x(1-x)$. Then the given sequence becomes[6] $a_1 = t$, $a_2 = f(t)$, $a_3 = f^2(t)$, and so on. Therefore, the problem asks to find the distinct roots of the equation $f^{1997}(t) = 0$.

[5]At least that was the situation at 1976. After this, the theory of iterations became increasingly known and popular and the IMO 1976 problem became also another textbook problem. We discuss the topic of iterations in Chapter 16.

[6]For more details on this idea, see Chapter 16.

First we notice that the image $f(x)$ of x will be in $[0,1]$ if $x \in [0,1]$. To have any roots, t must be in $[0,1]$. In such a case we can set $t = \sin^2\theta$, with $\theta \in [0, \pi/2]$. Then

$$f(t) = f(\sin^2\theta) = 4\sin^2\theta(1-\sin^2\theta) = (2\sin\theta\,\cos\theta)^2 = \sin^2(2\theta)\,.$$

And inductively

$$\begin{aligned} f^2(t) &= f(f(t)) = f(\sin^2(2\theta)) = \sin^2(4\theta)\,, \\ &\dots \\ f^n(t) &= \sin^2(2^n\theta)\,. \end{aligned}$$

The roots of $f^n(t) = 0$ are then those which satisfy $2^n\theta = k\pi$, $k \in \mathbb{Z}$ or, more precisely,

$$\theta = \frac{k\pi}{2^n}\,, \quad k = 0, 1, 2, \dots, 2^{n-1}\,.$$

For $n = 1997$, this gives $2^{1996} + 1$ distinct values of t. □

The function $f(x) = \lambda\, x(1-x)$ is known as the **logistic function** and plays an important role in the topic of *Chaos*. (See Section 16.7.) Sometimes it is also called the **population growth model**. This name is motivated by the following interpretation. Let $n = 1, 2, ..$ count the generations of a species and p_n be the population of the species at the n-th generation. If the population enjoys unlimited food supply and habitat, then it will grow according to the law $p_{n+1} = A\,p_n$ (geometric progression). However, as the population grows, stress develops over limited food supply and habitat. As a result, a fraction of the population dies. This 'removed' population is given by $-B\,p_n^2$. Therefore, the the population of each generation follows the law $p_{n+1} = Ap_n - Bp_n^2$. If we define $a_n = \lambda p_n$ and $\lambda = A/B$, then the last equation is written equivalently $a_{n+1} = \lambda a_n(1-a_n)$.

The following problem is an extension of Bolzano's theorem for discontinuous functions.

Problem 1.11 ([1], Problem E1336). *For the function $f : [0,1] \to \mathbb{R}$, $f(0) > 0$, $f(1) < 0$ and there exists a continuous function g such that $h = f + g$ is increasing. Prove that there exists a $\xi \in (0,1)$ such that $f(\xi) = 0$.*

Solution. Let

$$E = \{x \mid f(x) \geq 0\}\,.$$

This set is non-empty since $0 \in E$. It is also bounded since at most it can be $[0,1)$. Let $s \leq 1$ be its supremum. Since h is increasing, for any $x \in E$, $s \geq x$,

$$h(s) \geq h(x) = f(x) + g(x) \geq g(x)\,.$$

By taking the limit $x \to s$ of this inequality, since g is continuous, $h(s) \geq g(s)$, which, in turn, implies $f(s) \geq 0$.

Again from the increasing property of h, $h(s) \leq h(1)$ with $h(1) = g(1) + f(1) < g(1)$, $h(s) = g(s) + f(s) \geq g(s)$. Therefore $g(1) > h(1) \geq h(s) \geq g(s)$. Since g is continuous, it takes all values between $g(1)$ and $g(s)$. In particular it takes the value $h(s)$. That is, there exists a $\xi \geq s$ such that $g(\xi) = h(s)$.

Now we observe that

$$h(\xi) \geq h(s) \;\Rightarrow\; h(\xi) \geq g(\xi) \;\Rightarrow\; f(\xi) \geq 0 \,.$$

In other words, $\xi \in E$. However, by the definition of s, t cannot be greater than s; it must thus be $\xi = s$. Consequently, the equation $g(\xi) = h(s)$ gives $f(\xi) = 0$. □

Problem 1.12 (IMO 1968)**.** *Let f be a real-valued function defined for all real numbers x such that, for some positive a, the equation*

$$f(x+a) \;=\; \frac{1}{2} + \sqrt{f(x) - f(x)^2} \tag{1.2}$$

holds for all x.

(a) Prove that the function $f(x)$ is periodic.

(b) For $a = 1$ give an example of a non-constant function with the required properties.

Solution. (a) Setting $x = y + a$ in equation (1.2), we have:

$$\begin{aligned}
f(y+2a) &= \frac{1}{2} + \sqrt{f(y+a) - f(y+a)^2} \\
&= \frac{1}{2} + \sqrt{\frac{1}{2} + \sqrt{f(y) - f(y)^2} - \left(\frac{1}{2} + \sqrt{f(y) - f(y)^2}\right)^2} \\
&= \frac{1}{2} + \sqrt{\frac{1}{2} + \sqrt{f(y) - f(y)^2} - \frac{1}{4} - f(y) + f(y)^2 - \sqrt{f(y) - f(y)^2}} \\
&= \frac{1}{2} + \sqrt{\frac{1}{4} - f(y) + f(y)^2} \\
&= \frac{1}{2} + \sqrt{\left(\frac{1}{2} - f(y)\right)^2} \\
&= \frac{1}{2} + \left|\frac{1}{2} - f(y)\right| \,.
\end{aligned}$$

From this equation (or the given one) we see that $f(x) \geq 1/2$. Therefore, the absolute value found in the last equation is equal to $f(y) - 1/2$ and thus

$$f(y+2a) \;=\; f(y) \,.$$

the function $f(x)$ is periodic with period $2a$.

(b) The simplest function that is not constant is a function that takes two values. So, let

$$f(x) = \begin{cases} 1/2 & \text{if } x \in [0,1)\,, \\ 1 & \text{if } x \in [1,2)\,, \end{cases}$$

with the rest of the values values determined by periodicity, that is

$$f(x) = \begin{cases} 1/2 & \text{if } x \in [2n, 2n+1)\,, \\ 1 & \text{if } x \in [2n+1, 2n+2)\,, \end{cases}$$

for all $n \in \mathbb{Z}$. (See Figure 1.5 for its graph.)

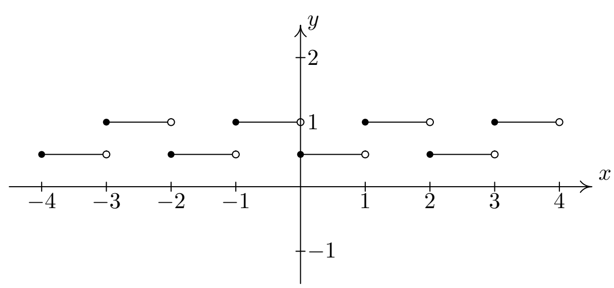

Figure 1.5: The graph of a discontinuous function that provides an example for the IMO 1968 problem.

Notice that the above function is discontinuous. If a continuous function is sought, one can use

$$f(x) = \frac{1}{2} + \frac{1}{2}\sin\left(\frac{\pi x}{2}\right), \quad x \in [0,2)\,,$$

with the rest of the values determined by periodicity. The result can be written nicely in the form

$$f(x) = \frac{1}{2} + \frac{1}{2}\left|\sin\left(\frac{\pi x}{2}\right)\right|\,,$$

for any $x \in \mathbb{R}$. (See Figure 1.6 for its graph.)

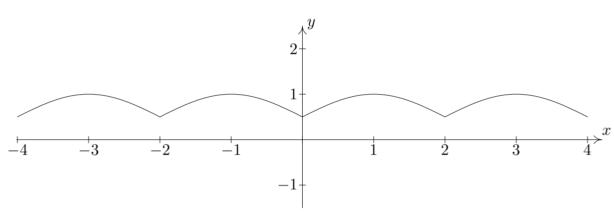

Figure 1.6: The graph of a continuous function that provides an example for the IMO 1968 problem.

However, the function is not differentiable at the points $x = n$, $n \in \mathbb{Z}$. We can give another example which is differentiable at these points too:

$$f(x) = \frac{1}{2} + \frac{1}{2} \sin^2\left(\frac{\pi x}{2}\right), \quad x \in \mathbb{R}.$$

(See Figure 1.7 for its graph.) □

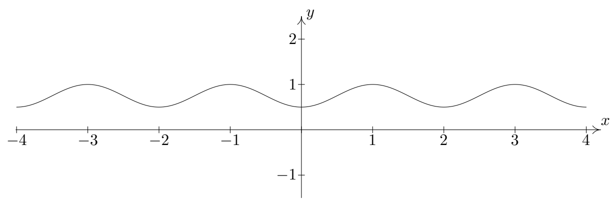

Figure 1.7: The graph of a differentiable function that provides an example for the IMO 1968 problem.

Problem 1.13 ([1], Problem 11233). *Show that for positive integer n and for $x \neq 0$,*

$$\frac{d^n}{dx^n}\left(x^{n-1} \sin \frac{1}{x}\right) = \frac{(-1)^n}{x^{n+1}} \sin\left(\frac{1}{x} + \frac{n\pi}{2}\right).$$

Solution. We can prove the given identity easily by induction. For $n = 1$

$$\begin{aligned} \frac{d}{dx} \sin \frac{1}{x} &= -\frac{1}{x^2} \cos \frac{1}{x} \\ &= -\frac{1}{x^2} \sin\left(\frac{1}{x} + \frac{\pi}{2}\right). \end{aligned}$$

That is, the identity is true.

Let it be true for some $n = k$:

$$\frac{d^k}{dx^k}\left(x^{k-1} \sin \frac{1}{x}\right) = \frac{(-1)^k}{x^{k+1}} \sin\left(\frac{1}{x} + \frac{k\pi}{2}\right).$$

Then, by the use of Theorem 1.11 and the above formula, we find

$$\begin{aligned} \frac{d^{k+1}}{dx^{k+1}}\left(x^k \sin \frac{1}{x}\right) &= \frac{d^{k+1}}{dx^{k+1}}\left(x\, x^{k-1} \sin \frac{1}{x}\right) \\ &= (k+1) \frac{d^k}{dx^k}\left(x^{k-1} \sin \frac{1}{x}\right) + x \frac{d^{k+1}}{dx^{k+1}}\left(x^{k-1} \sin \frac{1}{x}\right) \\ &= (k+1) \frac{(-1)^k}{x^{k+1}} \sin\left(\frac{1}{x} + \frac{k\pi}{2}\right) + x \frac{d}{dx} \frac{(-1)^k}{x^{k+1}} \sin\left(\frac{1}{x} + \frac{k\pi}{2}\right). \end{aligned}$$

The derivative of the product in the right hand side gives two terms one of which is opposite of the remaining term. So, finally

$$\begin{aligned}\frac{d^{k+1}}{dx^{k+1}}\left(x^k \sin\frac{1}{x}\right) &= \frac{(-1)^{k+1}}{x^{k+2}}\cos\left(\frac{1}{x}+\frac{k\pi}{2}\right)\\ &= \frac{(-1)^{k+1}}{x^{k+2}}\sin\left(\frac{1}{x}+\frac{(k+1)\pi}{2}\right).\end{aligned}$$

Therefore the identity is true for $n = k+1$ and consequently for all $n \in \mathbb{N}^*$. □

By the same technique, you may generalize the previous result to the following.

Problem 1.14. *If f is an n-times differentiable function, then*

$$\frac{d^n}{dx^n}\left[x^{n-1} f\left(\frac{1}{x}\right)\right] = \frac{(-1)^n}{x^{n+1}} f^{(n)}\left(\frac{1}{x}\right).$$

If this is too straightforward, a more challenging problem would be:

Problem 1.15. *If f is an n-times differentiable function, then find*

$$\frac{d^n}{dx^n}\left[x^m f\left(\frac{1}{x}\right)\right].$$

If you fail, the answer (but not the proof) is found in the January 2008 issue of [1].

Problem 1.16 ([1], Problem E3214)**.** *Let f be a real function with $n+1$ derivatives on $[a,b]$. Suppose $f^{(k)}(a) = f^{(k)}(b) = 0$ for $k = 0, 1, \ldots, n$. Prove that there is a number $\xi \in (a,b)$ such that $f^{(n+1)}(\xi) = f(\xi)$.*

Following the solution of R. Brooks ([1], vol. **96**, p. 740), we split the solution into two parts: a simple lemma (the case of $n = 0$) and the proof for a general n.

Lemma 1.1. *Let f be a differentiable function on $[a,b]$. Suppose $f(a) = f(b) = 0$. Prove that there is a number $\xi \in (a,b)$ such that $f'(\xi) = f(\xi)$.*

Proof. Consider the function

$$g(x) = e^{-x} f(x)$$

which satisfies the conditions of Rolle's theorem on $[a,b]$. Then there exists a $\xi \in (a,b)$ such that $g'(\xi) = 0$ or

$$e^{-\xi}\left(f(\xi) - f'(\xi)\right) = 0.$$

From this, it is obvious that $f'(\xi) = f(\xi)$. □

Using the above lemma we can now prove the statement in the general case.

Solution. Consider the function

$$g(x) = \sum_{k=0}^{n} f^{(k)}(x)$$

which satisfies the conditions of the previous lemma. Then there is a $\xi \in (a, b)$ such that $g'(\xi) = g(\xi)$ which is easily seen to give

$$f^{(n+1)}(\xi) = f(\xi) . \qquad \square$$

L.M. Levine has shown that the result holds true without requiring the vanishing of the derivatives at $x = b$. (See [1], vol. **96**, p. 740).

Part II

BASIC EQUATIONS

Chapter 2

Functional Relations Primer

2.1 The Notion of Functional Relations

Functional Equations

Functional equations are encountered routinely, even in introductory mathematics books. Some of the most common definitions of properties of functions are given using functional equations. For example, the definition of an **even function**,

$$f(x) \quad = \quad f(-x)\,, \quad \forall x\,,$$

the definition of an **odd function**,

$$f(x) \quad = \quad -f(-x)\,, \quad \forall x\,,$$

the definition of a **periodic function**,

$$\exists T \;:\quad f(x+T) \quad = \quad f(x)\,, \quad \forall x\,,$$

are all given in terms of functional equations.

So, intuitively, the notion of a functional equation seems to be straightforward. However, a general formal definition presents difficulties. The term "functional" refers to functions and thus includes any kind of equation one could possibly imagine: algebraic equations (i.e., those including just the functions), differential equations (i.e., those including the functions and their derivatives), integral equations (i.e., those including the functions and their integrals), difference equations (i.e., those including the functions evaluated at points differing by integers), integro-differential equations (i.e., those including the functions, their derivatives, and their integrals), and so on.

Example 2.1 (Truesdell equation)**.** Truesdell [43] has attempted to unify by a single theory the special functions of analysis using the differential-difference equation

$$\frac{\partial F(z,\alpha)}{\partial z} \;=\; F(z,\alpha+1)\,.$$

In this book we shall refrain from such advanced topics. □

Although we will meet some functional equations that contain derivatives and integrals, I will adopt the the standard convention which universally excludes[1] differential and integral equations. Some authors (for example, Saaty [33] and Hille [49]) tend to add to the list of exclusions the difference equations[2]. However, I reside with those authors (for example, Aczél [7]) who include the difference equations under the term "functional equations" as they are naturally embedded in the topic. And finally, the term excludes some advanced topics (such as operator functional equations) which, for the purposes of this book, I will pretend do not exist. What finally remains is the so-called *algebroid functional equations*. And yet, a solid general definition is still not easy to give. Trying to leave aside the difficulties so we can advance our study, we will start with the following 'working' definition:

Definition 2.1. An **algebroid functional equation** for the unknown function $f(x)$ is an equation of the form

$$F(x, f(x)) = 0\,, \quad \forall x \in D\,, \tag{2.1}$$

where F is a known function of two variables and D a given set.

Unfortunately definition (2.1) hides some important aspects of functional equations. The function F may contain parameters. Often these parameters enter in the form of variables that also take values in D. Other times the parameters can be written as the value of the function f itself. For example,

$$F(x, y, f(x), f(y)) = 0\,, \quad \forall x, y \in D\,, \tag{2.2}$$

is a functional equation of $f(x)$ with y and $f(y)$ appearing as 'parameters'. The fact that we permit parameters to enter in ways similar to this allows for a functional equation to have an endless number of possibilities. It is exactly this freedom that makes a rigid but widely applicable definition impossible. For our purposes, we shall assume that definition (2.1) is a shorthand notation for all such possibilities.

Comment. The reader should be warned that my terminology is not standard. Practitioners of the subject talk of functions of one or more variables and functional equations of one or more variables. That is, $f(x)$ is a *function of one variable*, but it can satisfy the *functional equation of two variables* $f(x+y) = f(x) + f(y)$. I have selected to reserve the term "variable" only for the functions, not the functional equations. Then I have selected the term "parameters" to describe the copies of the independent variable x and the dependent variable $f(x)$ that may appear in a given functional equation. What is usually referred to as a parameter is for me just a constant. Therefore the functional equation $f(f(x) + y) = f(x) + y + a$ is an equation for the function $f(x)$ of one variable and contains the parameter y and the constant a. The functional equation $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$ is an equation for the function $f(\vec{x}) = f(x_1, x_2, \ldots, x_n)$ of n variables and contains the $n+1$ parameters $\vec{y}, f(\vec{y})$.

[1] There is an extensive literature on such equations. If the reader is not already familiar with these topics, we recommend the following texts: [14],[25] for differential equations and [29],[42] for integral equations.

[2] There is also an extensive literature for difference equations. For example, see [18],[21].

My definition is motivated by the fact that I want to consider functional equations such as $f(f(x)) = a$ and $f(f(\vec{x})) = a$ of the same 'type'. In my definition, they are equations with "no parameters". According to the standard definition, one is an equation of one variable and the other is an equation of many variables.

In general, functional equations that contain "parameters" are easier to solve than equations that do not[3] . □

The set D is the domain of the function $f(x)$. If no domain is given explicitly, then it is implied that D should be taken to be all values of x for which the functional equation is truly an identity. The domain of a functional equation is important for the determination of the solutions; different domains lead to a different set of solutions. Also, given two different domains for an equation, nothing can be said a priori about the relation of the two sets of solutions.

Example 2.2. Let's consider the functional equation

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

This functional equation is one of the equations to be studied quite thoroughly in this book. If we seek functions $f : \mathbb{R}^* \to \mathbb{R}$, then the admissible solutions are $f(x) = 0$ or $f(x) = \ln|x|$. (See Chapter 5.) However, if we seek functions $f : \mathbb{R} \to \mathbb{R}$, if we set $y = 0$ in the equation, we find $f(0) = f(x) + f(0)$ that gives $f(x) \equiv 0$ as the unique solution. □

A function $f_p(x)$ that satisfies (2.1) is called a **particular solution**. Solving the functional equation (2.1) involves finding the set of all possible particular solutions. However, this set of solutions is strongly dependent on the properties imposed on the function $f(x)$. That is, one may search for measurable, invertible, bounded, monotonic, continuous, integrable, differentiable, etc., functions satisfying (2.1). Each condition on $f(x)$ will result, in general, in a different set of solutions.

Example 2.3. The identity function $f_p(x) = x$ is a particular solution of the equation $f(x+y) = f(x) + f(y)$ with $x, y \in \mathbb{R}$. The set of all continuous solutions is $\{f(x) = c\,x,\ c \in \mathbb{R}\}$. If we allow non-continuous solutions, then the set of solutions increases in size drastically. (See Sections 5.1 and 11.2.) □

Definition (2.1) may be generalized to a functional equation containing an unknown function of several variables,

$$F(x_1, x_2, \dots, x_n, f(x_1, x_2, \dots, x_n)) \ = \ 0 \ ,$$

[3] In [27], the term **iterative functional equations** is used for equations with no parameters. Although not all such equations contain iterates of functions, the term is motivated by the fact that iterates are essential for the methods and techniques used to obtain the solutions. (See Part V of this book.)

to a system of functional equations for several unknown functions $f_1(x), f_2(x), \ldots, f_N(x)$,

$$\begin{aligned} F_1(x, f_1(x), f_2(x), \ldots, f_N(x)) &= 0\,, \\ F_2(x, f_1(x), f_2(x), \ldots, f_N(x)) &= 0\,, \\ &\cdots\cdots \\ F_m(x, f_1(x), f_2(x), \ldots, f_N(x)) &= 0\,, \end{aligned}$$

or a combination of the above,

$$\begin{aligned} F_1(x_1, \ldots, x_n, f_1(x_1, \ldots, x_n), \ldots, f_N(x_1, \ldots, x_n)) &= 0\,, \\ F_2(x_1, \ldots, x_n, f_1(x_1, \ldots, x_n), \ldots, f_N(x_1, \ldots, x_n)) &= 0\,, \\ \cdots\cdots & \\ F_m(x_1, \ldots, x_n, f_1(x_1, \ldots, x_n), \ldots, f_N(x_1, \ldots, x_n)) &= 0\,. \end{aligned}$$

Of course, similar expressions with parameters are possible.

The Inverse Problem

So far we have presented the functional equation as the starting point for which one desires to know the solution. However, the inverse problem is very interesting: Given a function (or a set of functions), what functional equation (or equations) characterizes it (them)? If known, such a functional equation (or equations) can be used as the basis for the axiomatic definition of the function (or functions).

Example 2.4. One can define the well known trigonometric functions $\sin x$ and $\cos x$ (and thus all trigonometry) as follows.

Definition 2.2. Let the functions $S : \mathbb{R} \to \mathbb{R}$ and $C : \mathbb{R} \to \mathbb{R}$ satisfying the functional equation

$$C(x - y) = C(x)C(y) + S(x)S(y)\,, \quad \forall x, y \in \mathbb{R}\,, \tag{2.3}$$

and the following condition: There exists a positive number λ such that

$$C(0) = S(\lambda) = 1\,,$$

and

$$C(x) > 0\,, \quad S(x) > 0\,, \quad \forall x \in (0, \lambda)\,.$$

We call the solution thus obtained the **analytic sine** and **analytic cosine**.

Of course, one must prove that the above definition is meaningful; in other words, that there exists such pair of functions and it is unique. For example, notice that if

$$S(x + y) = S(x)C(y) + C(x)S(y)\,, \quad \forall x, y \in \mathbb{R}\,,$$

had be chosen in place of the functional equation written above, the lack of uniqueness is easily seen. The functions

$$C(x) = a^x \cos x\,, \quad S(x) = a^{x-\lambda} \sin x\,,$$

satisfy all requirements with $\lambda = \pi/2$ and any value of $a \in \mathbb{R}_+^*$. Therefore, an infinite number of solutions are obtained. We study the problem of characterization of the sine and cosine functions in Section 7.1. There, we will prove that the data given above for the analytic sine and cosine admit a unique solution, and this unique solution is given by $S(x) = \sin x$, $C(x) = \cos x$.

Leaving aside the issue of existence and uniqueness, one can easily show that the functions $S(x)$, $C(x)$ have all the properties of sine and cosine: $C(\lambda) = S(0) = 0$, $C(x)^2 + S(x)^2 = 1$, $|S(x)| \leq 1$, $|C(x)| \leq 1$, etc. (See Problem 2.14 on page 47.) □

Notice that, in the definition of analytic sine and analytic cosine, besides the functional equation, we used functional inequalities too. This is next section's topic of discussion.

Functional Relations

Functional equations may be generalized to functional relations in which case a function $f(x)$ is defined via a more complicated set of rules that contain more than equalities. Again, some of the most common definitions of properties of functions are given using functional relations. For example, the definition of an **increasing function**,

$$(f \text{ is } \uparrow) \Leftrightarrow \Big(\forall x < y\,, \quad f(x) \leq f(y)\Big)\,,$$

and the definition of a **decreasing function**,

$$(f \text{ is } \downarrow) \Leftrightarrow \Big(\forall x < y\,, \quad f(x) \geq f(y)\Big)\,,$$

are given in terms of functional inequalities. In general we can define:

Definition 2.3. A **functional relation** for the unknown function $f(x)$ is a relation of the form

$$F(x, f(x)) \sim G(x, f(x))\,, \quad \forall x \in D\,, \tag{2.4}$$

where F and G are known functions of two variables, D is a given set, and $\sim$ is some relation (usually an equivalence or ordering relation) defined among elements of D.

It is straightforward to generalize the above definition to more complicated cases that involve parameters, multiple relations, etc.

Problem 2.1 ([3], Problem 3039)**.** *Let a, b be fixed non-zero real numbers. Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f\left(x - \frac{b}{a}\right) + 2x \leq \frac{a}{b}x^2 + \frac{2b}{a} \leq f\left(x + \frac{b}{a}\right) - 2x, \quad \forall x \in \mathbb{R}\,.$$

Solution. Letting $y = x - b/a$ we see that the left inequality becomes

$$f(y) \leq \frac{a}{b}y^2 + \frac{2b}{a} .$$

Similarly, letting $y = x + b/a$ we see that the right inequality becomes

$$f(y) \geq \frac{a}{b}y^2 + \frac{2b}{a} .$$

Hence

$$f(y) = \frac{a}{b}y^2 + \frac{2b}{a} , \quad \forall y \in \mathbb{R} .$$

□

Some Additional Remarks

Having followed the previous discussion, the reader might find the remarks in this section beneficial and worth thinking of.

I have claimed that the 'definition' of functional equations given has excluded differential and integral functional equations according to the commonly used conventions. However, strictly speaking, this statement is not exactly right. Higher concepts in analysis, such as differentiation and integration, are based on simpler concepts of functions. Therefore, it is not surprising to discover differential and integral functional equations hidden inside our algebroid functional equations. Consider, for example, the following situation. For a continuous function of two variables $g(x, y)$ such that $g(x, 0) = 0$, let $f(x)$ be a continuous function of one variable satisfying the algebroid equation

$$f(x + y) - f(x) - y\, f(x) \ = \ y\, g(x, y) , \quad \forall x, y .$$

For $y = 0$, this equation appears to reduce to the trivial identity $0 = 0$. However, if we rewrite it as

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

at $y = 0$ it reduces to the differential equation

$$f'(x) \ = \ f(x) .$$

A comprehensive theory of functional equations that will be grouping them according to an underlying fundamental principle and which will encompass a structured way to solve them has proved to be extremely hard. To understand the reason, let's look at two similar equations: The so-callled **associativity functional equation**

$$f(x, f(y, z)) \ = \ f(f(x, y), z) ,$$

and the closely related equation

$$f(x, f(y, z)) \ = \ f(y, f(x, z)) ,$$

which I will not solve in this book (but the interested reader can find their solutions in Sections 6.2.1 and 6.2.3 of [7].) The two equations are identical in almost every property that we think matters: they both include only a single function of two variables; they both include just one additional parameter; they have identical 'structure'. Yet, the first equation is solved by

$$f(x, y) = g^{-1}(g(x) + g(y)) ,$$

and the second is solved by

$$f(x, y) = g^{-1}(h(x) + g(y)) .$$

The two answers are *very* different. In particular, the solution of the second equation includes an additional arbitrary function. So, besides hunches and intuition, it is not fully understood what is the underlying principle which differentiates two equations that appear identical in many respects. Therefore, creating a finer classification of functional equations and a general theory of how to solve them has been elusive. At the current time, we must rely on a classification that is based on empirical data. And such is the organization of this book.

2.2 Beginning Problems

As a leisurely introduction to solving functional equations, in this section we start with some simple problems which can be understood even by a student who is just learning the ideas of functions. As we progress with our subject in the next section and later chapters, we shall study progressively more difficult problems.

Problem 2.2. *Find all functions* $f : \mathbb{R} \to \mathbb{R}$ *such that* $2f(x+y)+6y^3 = f(x+2y)+x^3, \forall x, y \in \mathbb{R}$.

Solution. Direct: Substituting $y = 0$ in the functional equation, we get $2f(x) = f(x) + x^3$ or $f(x) = x^3$.

Inverse: For $f(x) = x^3$, the left hand of the given equation equals

$$2x^3 + 6x^2y + 6xy^2 + 8y^3 ,$$

while the right hand side equals

$$2x^3 + 6x^2y + 12xy^2 + 8y^3 ,$$

The two terms are not equal and therefore the equation has no solution. □

The previous problem points out that, when solving a functional equation, we must check that the derived functions indeed satisfy the given equation (inverse). However, in the following, we systematically omit this step whenever everything works out fine.

Problem 2.3. *Let* $f : \mathbb{R}_+^* \to \mathbb{R}$ *such that*

$$f(x^2) \; = \; \frac{1}{x} \, , \quad \forall x \in \mathbb{R}_+^* \, .$$

Find f.

Solution. In the defining equation we set $y = x^2$. Then

$$f(y) \; = \; \frac{1}{\sqrt{y}} \, . \qquad \square$$

Problem 2.4. *Let* $f : \mathbb{R}^* \to \mathbb{R}$ *such that*

$$f(1 + \frac{1}{x}) \; = \; x^2 + \frac{1}{x^2} \, , \quad \forall x \in \mathbb{R}^* \, .$$

Find f.

Solution. In the defining equation we set $y = 1 + 1/x = (x+1)/x$:

$$f(y) \; = \; \frac{1}{(y-1)^2} + (y-1)^2 \, . \qquad \square$$

Problem 2.5. *Find a particular solution* $f : \mathbb{R}^* \to \mathbb{R}$ *of the functional equation:*

$$f(x) - f(x+1) \; = \; f(x)\, f(x+1) \, , \quad \forall x \in \mathbb{R}^* \, .$$

Solution. The function $f_p(x) = 1/x$ satisfies the equation. Indeed

$$\begin{aligned} f_p(x) - f_p(x+1) \; &= \; \frac{1}{x} - \frac{1}{x+1} \\ &= \; \frac{1}{x(x+1)} \\ &= \; f_p(x)\, f_p(x+1) \, , \quad x \neq 0 \, , -1 \, . \end{aligned} \qquad \square$$

Problem 2.6. *(a) Show that the function*

$$f_p(x) \;=\; a\,\cos^{-1} x\,,\quad 0 \le x \le 1\,,$$

where a is a real constant, satisfies the functional equation

$$f(xy - \sqrt{(1-x^2)(1-y^2)}) \;=\; f(x) + f(y)\,.$$

(b) Find a particular solution $f : [1,\infty) \to \mathbb{R}$ *of the functional equation*

$$f(xy - \sqrt{(x^2-1)(y^2-1)}) \;=\; f(x) + f(y)\,.$$

(c) Find a particular solution $f : \mathbb{R} \to \mathbb{R}$ *of the functional equation*

$$f(x\sqrt{1+y^2} + y\sqrt{1+x^2}) \;=\; f(x) + f(y)\,.$$

(d) Find a particular solution $f : [-1,1] \to \mathbb{R}$ *of the functional equation*

$$f(\frac{x+y}{1-xy}) \;=\; f(x) + f(y)\,.$$

(e) Find a particular solution $f : \mathbb{R} \to \mathbb{R}$ *of the functional equation*

$$f(\frac{x+y}{1+xy}) \;=\; f(x) + f(y)\,.$$

Solution. (a) Let $X = \cos^{-1} x$ and $Y = \cos^{-1} y$. Then $x = \cos X$ and $y = \cos Y$ and

$$f(x) + f(y) \;=\; a(X+Y)\,.$$

On the other hand

$$\begin{aligned} xy - \sqrt{(1-x^2)(1-y^2)} \;&=\; \cos X \cos Y - \sqrt{\sin^2 X \sin^2 Y} \\ &=\; \cos X \cos Y - \sin X \sin Y \\ &=\; \cos(X+Y)\,. \end{aligned}$$

Therefore

$$f(xy - \sqrt{(1-x^2)(1-y^2)}) \;=\; a\,\cos^{-1}(\cos(X+Y)) \;=\; a\,(X+Y)\,,$$

from which the relation sought follows immediately.

(b)–(e) Similarly, we see that these functional equations are satisfied by the functions $a\cosh^{-1} x$, $a\sinh^{-1} x$, $a\tan^{-1} x$, and $a\tanh^{-1} x$, respectively. $\square$

Problem 2.7. *Find all functions* $f : \mathbb{R} \to \mathbb{R}$ *such that*

$$f(x+y) \;=\; r\,y + f(x)\,, \quad \forall x \in \mathbb{R}\,, \quad \forall y \in \mathbb{R}\,.$$

Solution. Substituting $x = 0$ in the defining functional equation, we find

$$f(y) \;=\; r\,y + c\,,$$

where $c = f(0)$. □

A variation of the previous problem is as follows.

Problem 2.8. *Find all functions* $f : \mathbb{R} \to \mathbb{R}$ *such that*

$$f(x+y) \;=\; r\,y + f(x)\,, \quad \forall x \in \mathbb{R}\,, \quad \forall y \in \mathbb{R}_+^*\,.$$

Solution. Substituting $x = 0$ in the defining functional equation, we find

$$f(y) \;=\; r\,y + c\,, \quad y > 0\,.$$

where $c = f(0)$. We then substitute $x = -y$, $y > 0$ to find

$$f(0) \;=\; r\,y + f(-y) \;\Rightarrow\; f(-y) \;=\; -\,r\,y + c\,.$$

Therefore

$$f(x) \;=\; r\,x + c\,, \quad \forall x\,.$$ □

The last two problems imply the following result:

Corollary. *The only periodic function* $f(x)$

$$f(x+T) \;=\; f(x)\,, \quad \forall x \in \mathbb{R}$$

with periods all positive real numbers in the unit interval (0,1] is the constant function.

Proof. Since $T = 1$ is a period, $f(x+1) = f(x)$. Then substituting $x+1$ in the place of x in the functional equation gives

$$f(x+T+1) \;=\; f(x+1) \;=\; f(x)\,.$$

Therefore, if T is a period, $T+1$ is a period too. Inductively, $T+n$, for any natural number is a period too. Substituting $x - T$ into the original equation gives that $f(x) = f(x-T)$, that is, if T is a period, $-T$ is a period too. So, if all points (0,1] are periods of $f(x)$, all points y of $\mathbb{R}$ are periods. and the result follows by setting $r = 0$ in the preceding problems. □

Problem 2.9. *Find all functions* $f : \mathbb{R} \to \mathbb{R}$ *such that*

$$f(xy) = y^r f(x) , \quad \forall x, y \in \mathbb{R} ,$$

where we have assumed that the constant r *is such that* y^r *is well defined for all real* y*.*

Solution. Setting $x = 1$ in the defining equation we find $f(y) = c\, y^r$, with $c = f(1)$. Inversely, the function $f(x) = c\, x^r$ satisfies the given functional equation. □

Problem 2.10. *Find all functions* $f : \mathbb{R} \to \mathbb{R}$ *such that*

$$f(xy) = y^r f(x) , \quad \forall x \in \mathbb{R} , \quad \forall y \in \mathbb{R}_+^* .$$

Solution. Setting $x = 0$ in the defining equation we find

$$f(0)\,(1 - y^r) = 0 .$$

If $r \neq 0$, since this equation must be true for all positive y, we must have $f(0) = 0$. If $r = 0$, then the value $f(0)$ is undetermined.

Then we substitute $x = 1$, $x = -1$ in the defining equation to find

$$\begin{aligned} f(y) &= a\, y^r , \quad y > 0 , \\ f(-y) &= b\, y^r , \quad y > 0 , \end{aligned}$$

where we have set $a = f(1)$ and $b = f(-1)$. From the above results we see that for $r \neq 0$

$$f(x) = \begin{cases} a\, x^r , & \text{if } x > 0 , \\ 0 , & \text{if } x = 0 , \\ b\, |x|^r , & \text{if } x < 0 , \end{cases}$$

while for $r = 0$

$$f(x) = \begin{cases} a , & \text{if } x > 0 , \\ c , & \text{if } x = 0 , \\ b , & \text{if } x < 0 , \end{cases}$$

Inversely, it is a simple calculation to show that the above functions satisfy the given functional equation. □

Problem 2.11. *Find the function $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(\lambda x, \lambda y) \;=\; \lambda^r f(x,y)\,, \quad \forall \lambda, x, y\,,$$

and r a fixed number.

Solution. Setting $\lambda = 1/x$, $x \neq 0$, we find

$$f(x,y) \;=\; x^r f\left(1, \frac{y}{x}\right)\,.$$

That is, if $g(y) = f(1,y)$ is any real-valued function, then it gives a solution

$$f(x,y) \;=\; x^r g\left(\frac{y}{x}\right)\,,$$

of the given equation. □

Comment. It is straightforward to generalize the above problem to one in n variables: The equation

$$f(\lambda x_1, \lambda x_2, \ldots, \lambda x_n) \;=\; \lambda^r f(x_1, x_2, \ldots, x_n)\,, \quad \forall \lambda, x_1, x_2, \ldots, x_n\,, \tag{2.5}$$

is solved by

$$f(x_1, x_2, \ldots, x_n) \;=\; x_1^r\, g\left(\frac{x_2}{x_1}, \ldots, \frac{x_n}{x_1}\right)\,,$$

where g is any function in $n-1$ variables. □

Comment. The functional equation (2.5) is known as the **Euler equation** and a function satisfying it is called **homogeneous of degree** r. □

Problem 2.12. *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that*
(a) $f(\lambda x) \;=\; f(x)\,, \quad \forall x \in \mathbb{R}\,,$
where λ is a real constant not ± 1 and
(b) f is continuous at $x = 0$.

Solution. We shall study three different cases.

(i) If $\lambda = 0$ then it is immediate that $f(x) = f(0)$.

(ii) If $|\lambda| < 1$, then by repeating applications of the defining functional equation

$$f(x) \;=\; f(\lambda x) \;=\; f(\lambda^2 x) \;=\; \cdots \;=\; f(\lambda^n x)\,,$$

for any natural number n. Then, by the continuity of f at $x = 0$:

$$f(0) \;=\; f(\lim_{n\to\infty} \lambda^n x) \;=\; \lim_{n\to\infty} f(\lambda^n x) \;=\; \lim_{n\to\infty} f(x) \;=\; f(x)\,.$$

(iii) If $|\lambda| > 1$, we define $\mu = 1/\lambda$ and in the defining equation we substitute $x = \mu y$: $f(y) = f(\mu y)$. Then

$$f(y) = f(\mu y) = f(\mu^2 y) = \cdots = f(\mu^n y)\,,$$

for any natural number n. Using the the continuity of f at $x = 0$, we conclude again that

$$f(x) = f(0)\,.$$

Inversely, the constant function $f(x) = c$ satisfies the given conditions. □

Question. *Let now $h(x)$ be any continuous periodic function with period T, i.e. $h(x + T) = h(x)$, $\forall x \in \mathbb{R}$. Then define*

$$f(x) = h(a + T \log_\lambda x)\,,$$

where $a \in \mathbb{R}$ and $\lambda \in \mathbb{R}_+^* \setminus \{1\}$. Show that

$$f(\lambda x) = f(x)\,, \quad \forall x \in \mathbb{R}_+^*\,.$$

However, f is not identically zero. Why?

Answer. Although the function $f(x)$ satisfies the functional equation $f(\lambda x) = f(x)$, it does not satisfy the second condition: continuity at $x = 0$. This condition was essential in order to prove that the function is constant. If we eliminate this condition, then non-constant solutions of the functional equation are possible (as demonstrated explicitly by the given function).

Problem 2.13. *Find the function $f : \mathbb{R} \to \mathbb{R}$ such that*

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

Solution. The identically vanishing function $f(x) \equiv 0$ is a solution.

We shall search for non-identically vanishing solutions. That is, there exists a $x_0 \in \mathbb{R}$ such that $f(x_0) \neq 0$. Setting $x = x_0$ and $y = 0$ in the defining solution, we find $f(x_0) = f(x_0)\,f(0)$ and therefore $f(0) = 1$. Setting now $y = x$, we find $f(0) = f(x)^2$ or $f(x) \equiv \pm 1$. The solution $f(x) \equiv -1$ is inconsistent with the defining relation. Hence $f(x) \equiv 1$. □

Problem 2.14. *Consider the functions of analytic sine and analytic cosine as defined on page 38. Prove that:*

(a) $S(0) = C(\lambda) = 0$.

(b) $C(x)^2 + S(x)^2 = 1$.

(c) The two functions $C(x)$, $S(x)$ are bounded. In particular, $|C(x)| \leq 1$ and $|S(x)| \leq 1$.

(d) $C(\lambda - x) = S(x)$ and $S(\lambda - x) = C(x)$.

(e) The functions $C(x)$, $S(x)$ are periodic.

(f) $S(x + y) = S(x)C(y) + S(y)C(x)$.

(g) $C(-x) = C(x)$ and $S(-x) = -S(x)$.

(h) For the functions $C(x)$, $S(x)$ all known trigonometric identities for $C(x \pm y)$ and $S(x \pm y)$ and their corollaries are valid.

(i) The functions $C(x)$, $S(x)$ are continuous in $\mathbb{R}$.

Solution. (a) In the defining equation (2.3), we set $x = y = 0$ to find

$$C(0) \;=\; C(0)^2 + S(0)^2 \,,$$

which, when combined with $C(0) = 1$, gives $S(0) = 0$. Similarly, if in the defining equation (2.3), we set $x = y = \lambda$, we find $C(\lambda) = 0$.

(b) This identity is now immediate if, in the defining equation (2.3), we set $x = y$.

(c) If we write the previous result in the form $|C(x)|^2 + |S(x)|^2 = 1$, we conclude that $|C(x)|^2 \leq 1$, $|S(x)|^2 \leq 1$, or equivalently $|C(x)| \leq 1$, $|S(x)| \leq 1$.

(d) In the defining equation (2.3), we set $x = \lambda$ to find

$$C(\lambda - y) \;=\; C(\lambda)C(y) + S(\lambda)S(y) \;\Rightarrow\; C(\lambda - y) \;=\; S(y) \,.$$

In the last equation, set $y = \lambda - x$ to find $C(x) = S(\lambda - x)$ too.

(e) In the result (d), place $x - \lambda$ in place of x:

$$\begin{aligned} C(2\lambda - x) &= S(x - \lambda) \;=\; -\,S(\lambda - x) \;=\; -\,C(x) \,, \\ S(2\lambda - x) &= C(x - \lambda) \;=\; +\,C(\lambda - x) \;=\; +\,S(x) \,. \end{aligned}$$

Substituting once more $-x - 2\lambda$ for x we arrive at the result sought:

$$\begin{aligned} C(4\lambda + x) &= -C(-x - 2\lambda) \;=\; -\,C(2\lambda + x) \;=\; C(-x) \;=\; C(x) \,, \\ S(4\lambda + x) &= S(-x - 2\lambda) \;=\; -\,S(2\lambda + x) \;=\; -\,S(-x) \;=\; S(x) \,. \end{aligned}$$

(f) Using the result (d), we can write:

$$\begin{aligned} S(x + y) &= C(\lambda - x - y) \;=\; C((\lambda - x) - y) \\ &= C(\lambda - x)C(y) + S(\lambda - x)S(y) \\ &= S(x)C(y) + C(x)S(y) \,. \end{aligned}$$

(g) If, in the defining equation (2.3), we set $x = 0$ then we immediately see that the analytic cosine is even: $C(-y) = C(y)$. Now, in the equation we proved in the previous part (f), we set $y = -x$:

$$S(0) \;=\; S(x)C(-x) + S(-x)C(x) \;\Rightarrow\; 0 \;=\; C(x)\,[S(x) + S(-x)] \,.$$

If for any x, $C(x) \neq 0$, then $S(x) = -S(-x)$. This is the case, for example, for any $x \in (0, \lambda)$. However, for any x' satisfying $C(x') = 0$, no information is obtained for this equation. However, in this case, the result (b) requires that $S(x') \pm 1$. Let's take an arbitrary $y \in (0, \lambda)$. Then

$$C(x' + y) \;=\; C(x' - (-y)) \;=\; C(x')C(y) + S(x')S(-y) \;=\; \pm\,S(y) \,.$$

This result implies that $C(x' + y) \neq 0$ and therefore

$$S(x' + y) \;=\; -\,S(-x' - y) \,.$$

Using part (e), this equation is equivalently written

$$S(x')C(y) \;=\; -\,S(-x')C(y)\;,$$

or $S(x') = -S(-x')$.

(h) This part is a straightforward application of the formulæ found so far. For example, setting $x = -y$ in the defining equation (2.3), we find

$$\begin{aligned} C(2x) &= C(x)^2 - S(x)^2 \\ &= 2C(x)^2 - 1 \\ &= 1 - 2S(x)^2\;. \end{aligned}$$

This can also be written as

$$C(x) \;=\; \pm\sqrt{\frac{1 + C(2x)}{2}}\;,$$

with the plus sign chosen if $x \in (0, \lambda)$. One can continue similarly.

Notice that one can show for these results that $C(x)$ is strictly decreasing in the interval $[0, 2\lambda]$ and strictly increasing in the interval $[2\lambda, 4\lambda]$. $S(x)$ is strictly increasing in the interval $[-\lambda, \lambda]$ and strictly decreasing in the interval $[\lambda, 3\lambda]$. The reader may want to fill the details.

(i) Since $0 \le C(x) \le 1$ for any $x \in (0, \lambda)$, the limit $\lim\limits_{x\to 0^+} C(x)$ exists and it is a number in the interval $[0, 1]$. Let's indicate it by ℓ. Then from the equation $C(2x) = 2C(x)^2 - 1$ we find that ℓ satisfies the equation $\ell = 2\ell^2 - 1$. The positive root of this equation is $\ell = 1$. Since $C(x)$ is even, the root from the other side exists and is also 1. Therefore

$$\lim_{x\to 0} C(x) \;=\; 1 \;=\; C(0)\;,$$

that is, $C(x)$ is continuous at $x = 0$.

Using the equation $C(2x) = 1 - 2S(x)^2$, we also see that $S(x)$ is continuous at $x = 0$ too:

$$\lim_{x\to 0} S(x) \;=\; 0 \;=\; S(0)\;,$$

Now using the equation

$$C(x + h) \;=\; C(x)C(h) - S(x)S(h)\;,$$

and taking the limit $h \to 0$, we find

$$\lim_{h\to 0} C(x + h) \;=\; C(x)\lim_{h\to 0} C(h) - S(x)\lim_{h\to 0} S(h)\;,$$

or

$$\lim_{h\to 0} C(x + h) \;=\; C(x)\;.$$

The function $C(x)$, and thus $S(x)$ are continuous at all points. □

Chapter 3

Equations for Arithmetic Functions

3.1 The Notion of Difference Equations

An **arithmetic function** is a function $a(n)$ from the set of natural numbers $\mathbb{N}$ (or one of its subsets) to $\mathbb{R}$ (or $\mathbb{C}$, or any of their subsets). This terminology is common in number theory. In calculus, one usually speaks of a *sequence* and often writes a_n as an equivalent notation to $a(n)$. We will use both notations at convenience.

A **difference functional equation** is an equation that relates a number of terms of the sequence:

$$f(a_{n+1}, a_n, a_{n-1}, \ldots, a_0, n) = 0 \,,$$

where f is some given function. If the difference equation can be solved for a_{n+1}, say

$$a_{n+1} = g(a_n, a_{n-1}, \ldots, a_0, n) \,,$$

we speak of a **recursion relation**.

Example 3.1. The equation

$$a_{n+3}^3 + a_{n+2} \ln a_{n+3} + a_{n+1} \tan a_{n+3} - \frac{a_n}{a_{n+2}} = 0$$

is a difference equation that relates any term of the sequence a_n to the three preceding terms. It is not easy however to solve for a_{n+3}.

The equations

$$a_{n+1} = a_n + \omega \,, \tag{3.1}$$

$$a_{n+1} = \lambda\, a_n \,, \tag{3.2}$$

$$a_{n+1} = a_n + a_{n-1} \,, \tag{3.3}$$

are some of the most known recursion relations. Equation (3.1) defines an **arithmetic progression** with **step** ω:

$$a_0, \quad a_0 + \omega, \quad a_0 + 2\omega, \quad a_0 + 3\omega, \ldots$$

Equation (3.2) defines a **geometric progression** with **ratio** λ:

$$a_0, \quad a_0\lambda, \quad a_0\lambda^2, \quad a_0\lambda^3, \ldots$$

Equation (3.3) defines a sequence that depends on the first two terms a_0 and a_1. If $a_0 = a_1 = 1$, then the so-called **Fibonnacci sequence** is found:

$$1, \quad 1, \quad 2, \quad 3, \quad 5, \quad 8, \quad 13, \cdots .$$

The members of the sequence are called the **Fibonnacci numbers**. Similarly, if $a_0 = 1$, $a_1 = 3$, then the so-called **Lucas sequence** is found:

$$1, \quad 3, \quad 4, \quad 7, \quad 11, \quad 18, \quad 29, \cdots .$$

The members of the sequence are called the **Lucas numbers**. □

Exercise 3.1. *(a) If $\{a_n\}$ is an arithmetic progression prove that*

$$\sum_{k=1}^{n} a_k = \frac{a_1 + a_n}{2} n .$$

(b) If $\{a_n\}$ is a geometric progression with ratio λ prove that

$$\sum_{k=1}^{n} a_k = \frac{a_n\lambda - a_1}{\lambda - 1} .$$

If $|\lambda| < 1$, then

$$\sum_{k=1}^{n} a_k = \frac{a_1}{1 - \lambda} .$$ □

Formal Definition

We can make the definition of a difference equation a bit more formal by the introduction of two *operators* (that is, objects that operate on functions and produce new functions).

Given a function $a(n)$, the **translation operator** (or **shift operator**) $\hat{T}$ is defined by

$$\hat{T}a(n) = a(n+1) .$$

Applying the operator successively m times, $\hat{T}^m = \hat{T}\,\hat{T}^{m-1}$, $m > 1$, we can easily see that

$$\hat{T}^m a(n) = a(n+m) .$$

The **difference operator** $\hat{\Delta}$ is defined by

$$\hat{\Delta}a(n) = a(n+1) - a(n) . \tag{3.4}$$

We can also apply this operator many times, $\hat{\Delta}^m = \hat{\Delta}\,\hat{\Delta}^{m-1}$, $m > 1$. For $m = 2$:

$$\begin{aligned}
\hat{\Delta}^2 a(n) &= \hat{\Delta}\Big(\hat{\Delta}a(n)\Big) \\
&= \hat{\Delta}\Big(a(n+1) - a(n)\Big) \\
&= [a(n+2) - a(n+1)] - [a(n+1) - a(n)] \\
&= a(n+2) - 2a(n+1) + a(n)\,.
\end{aligned}$$

This last equation may be written as

$$\hat{\Delta}^2 a(n) \;=\; (\hat{T}^2 - 2\hat{T} + \hat{I})a(n) \;=\; (\hat{T} - \hat{I})^2 a(n)\,,$$

where we introduced the **identity operator**

$$\hat{I}a(n) \;=\; a(n)\,.$$

Inductively we can show that

$$\hat{\Delta}^m \;=\; (\hat{T} - \hat{I})^m \;=\; \sum_{k=0}^{m} \binom{m}{k} (-1)^{m-k} \hat{T}^k\,.$$

Note that the definition of the difference operator $\hat{\Delta}$ implies that

$$\hat{\Delta} \;=\; \hat{T} - \hat{I}$$

or

$$\hat{T} \;=\; \hat{\Delta} + \hat{I}\,.$$

Consequently,

$$\hat{T}^m \;=\; (\hat{\Delta} + \hat{I})^m \;=\; \sum_{k=0}^{m} \binom{m}{k} \hat{\Delta}^k\,.$$

The last equation implies that any term $a(n+m)$ of the sequence $\{a(n)\}$ can be written in terms of the m differences $\hat{\Delta}^k a(n)$, $k = 1, 2, \ldots m$. We have thus arrived at an alternative definition of difference equation:

Definition 3.1. An equation relating the value of the function $a(n)$ with one or more of its differences $\hat{\Delta}^k a(n)$, $k = 1, 2, \ldots,$ is called a **difference equation**.

Example 3.2. Some examples of difference equations are

$$\begin{aligned}
\hat{\Delta}a(n) + 5a(n) &= 0\,, \\
\hat{\Delta}^2 a(n) + \hat{\Delta}a(n) + a(n) &= n^2 + 1\,, \\
\Big(\hat{\Delta}a(n)\Big)^2 + a_n^2 &= 1\,, \\
a(n)\,\hat{\Delta}^3 a(n) &= \cos(3n)\,.
\end{aligned}$$

If the differences are substituted for their values, we can rewrite these equations as

$$\begin{aligned} a_{n+1} + 4a_n &= 0\,, \\ a_{n+2} - a_{n+1} - a_n &= n^2 + 1\,, \\ a_{n+1}^2 - 2a_{n+1}\, a_n + 2a_n^2 &= 1\,, \\ a_n\, a_{n+3} - 3a_n\, a_{n+2} + 3a_n\, a_{n+1} - a_n^2 &= \cos(3n)\,. \end{aligned}$$ □

3.2 Multiplicative Functions

Definition 3.2. An arithmetic function $f(n)$ is called **multiplicative** if f is not identically zero and

$$f(m\,n) = f(m)\, f(n)\,, \tag{3.5}$$

for all $m, n \in \mathbb{N}^*$ which are coprime. It is called **completely multiplicative** if

$$f(m\,n) = f(m)\, f(n)\,, \quad \forall n, m\,. \tag{3.6}$$

Example 3.3. The **Möbius function** $\mu(n)$ is defined as follows: If $n = 1$, then $\mu(1) = 1$. If $n \neq 1$, let $p_1, p_2, \ldots, p_k$ be the prime factors of n, $n = p_1^{a_1} p_2^{a_2} \ldots p_k^{a_k}$. Then

$$\mu(n) = \begin{cases} (-1)^k\,, & \text{if } a_1 = a_2 = \cdots = a_k = 1\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$ □

It is straightforward to see that the Möbius function is multiplicative but not completely multiplicative.

We can easily prove the following statements for multiplicative functions:

Theorem 3.1. *If f, g are two (completely) multiplicative functions, then fg and f/g are also (completely) multiplicative functions.*

Theorem 3.2. *If f is multiplicative, then $f(1) = 1$.*

Theorem 3.3. *Let f be an arithmetic function such that $f(1) = 1$. Then f will be multiplicative iff*

$$f(p_1^{a_1} p_2^{a_2} \ldots p_k^{a_k}) = f(p_1^{a_1})\, f(p_2^{a_2}) \ldots f(p_k^{a_k})\,,$$

for all different prime numbers p_i and all integers $a_i \geq 1$.

Theorem 3.4. *A multiplicative function f is completely multiplicative iff*

$$f(p^a) = f(p)^a\,,$$

for all prime numbers p and integers $a \geq 1$.

The proofs of the above theorems are left to the reader as easy exercises.

3.3 Linear Difference Equations

A difference equation of the form

$$a_{n+k} + p_1(n)\,a_{n+k-1} + \cdot + p_k(n)\,a_n \;=\; g(n)\;, \tag{3.7}$$

with $p_i(n)$, $i = 1, 2, \ldots, k$ given functions, is called a **linear difference equation of k-th order**. Furthermore, it is called **homogeneous** if $g(n) = 0$ and **non-homogeneous** if $g(n) \neq 0$. In physics, this function is related to a driving force that tries to impose (drive/force) a particular behavior on the system. For this reason, $g(n)$ is sometimes called the **forcing term**. The term *linear* means that if the sequences b_n and c_n are solutions of the homogeneous equation, then so is the linear combination $\lambda b_n + \mu c_n$ for any numbers λ, μ (as can be easily verified).

In order to solve a k-th order difference equation, initial data must be given. The initial data are usually the first k terms:

$$a_0, a_1, \ldots, a_{k-1} = \text{given}\;. \tag{3.8}$$

The equation (3.7) together with the initial data (3.8) constitutes an initial value problem. We have the following theorem:

Theorem 3.5. *The initial value problem (3.7), (3.8) has a unique solution.*

Proof. Given (3.8), we can find a_k from (3.7), then a_{k+1}; then we can find a_{k+2}, and so on. Therefore, by induction, any a_n is determined uniquely. □

Obviously, the above theorem not only guarantees that a solution exists, but it also determines how it can be constructed. However, such an inductive approach is very inefficient, requiring a lot of work and time in many cases. Therefore, a more concise approach is desirable; such an approach is presented in the remainder of the section.

Theory of Solutions

In this subsection we present general statements about the solutions of linear difference equations. In the next subsection we then describe how to find the solutions of such difference equations building on the results of this subsection.

Definition 3.3. The functions $f_1(n), f_2(n), \ldots, f_r(n)$ are **linearly dependent** for $n \geq n_0$ if there are constants $\alpha_1, \alpha_2, \ldots, \alpha_r$ such that

$$\alpha_1 f_1(n) + \alpha_2 f_2(n) + \cdots \alpha_r f_r(n) \;=\; 0\;,$$

for all $n \geq n_0$ and

$$|\alpha_1| + |\alpha_2| + \cdots + |\alpha_r| \neq 0\;.$$

A set of functions that are not linearly dependent are called **linearly independent**.

Exercise 3.2. *Show that 3^n, $n3^n$, $n^2 3^n$ are linearly independent for $n \geq 1$.*

Solution. Let $\alpha_1, \alpha_2, \alpha_3$ be such that

$$\alpha_1 3^n + \alpha_2 n 3^n + \alpha_3 n^2 3^n = 0 .$$

We then get

$$\alpha_1 + \alpha_2 n + \alpha_3 n^2 = 0 .$$

For $n = 1, 2, 3$ we get the system of equations:

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= 0 , \\ \alpha_1 + 2\alpha_2 + 4\alpha_3 &= 0 , \\ \alpha_1 + 3\alpha_2 + 9\alpha_3 &= 0 , \end{aligned}$$

with unique solution $\alpha_1 = \alpha_2 = \alpha_3 = 0$. ◻

Definition 3.4. A set of k linearly independent solutions for the difference equation (3.7) is called a **fundamental set** of solutions.

Given a set of solutions $x_1(n), x_2(n), \ldots, x_r(n)$ of (3.7), there is a straightforward method to check for linear independence, as explained below.

Definition 3.5. The **Casoratian** of the solutions $x_1(n), x_2(n), \ldots, x_r(n)$ is defined by:

$$C(n) = \det \begin{bmatrix} x_1(n) & x_2(n) & \cdots & x_r(n) \\ x_1(n+1) & x_2(n+1) & \cdots & x_r(n+1) \\ \vdots & \vdots & & \vdots \\ x_1(n+r-1) & x_2(n+r-1) & \cdots & x_r(n+r-1) \end{bmatrix} .$$

Theorem 3.6. *The set of k solutions $x_1(n), x_2(n), \ldots, x_k(n)$ are linearly independent iff $C(n) \neq 0$.*

Proof. Suppose that there exist constants $\alpha_1, \alpha_2, \ldots \alpha_k$ such that

$$\alpha_1 x_1(n) + \alpha_2 x_2(n) + \cdots + \alpha_k x_k(n) = 0 .$$

Then it is clear that the equations

$$\begin{aligned} a_1 x_1(n+1) + a_2 x_2(n+1) + \cdots + a_k x_k(n+1) &= 0 , \\ a_1 x_1(n+2) + a_2 x_2(n+2) + \cdots + a_k x_k(n+2) &= 0 , \\ \cdots \\ a_1 x_1(n+k-1) + a_2 x_2(n+k-1) + \cdots + a_k x_k(n+k-1) &= 0 , \end{aligned}$$

must also be satisfied. The set of the above equations can be written in a matrix form as:

$$\begin{bmatrix} x_1(n) & x_2(n) & \cdots & x_k(n) \\ x_1(n+1) & x_2(n+1) & \cdots & x_k(n+1) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(n+k-1) & x_2(n+k-1) & \cdots & x_k(n+k-1) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} .$$

We know that the solution $a_1 = a_2 = \cdots = a_k = 0$ is a unique solution if and only if the determinant of the above $k \times k$ matrix is nonzero, or $C(n) \neq 0$. Thus the aforementioned set of equations are linearly independent if and only if $C(n) \neq 0$. □

Example 3.4. The equation

$$a_{n+3} - 7a_{n+1} + 6a_n \;=\; 0$$

has the solutions

$$x_1(n) = 1, \quad x_2(n) = (-3)^n, \quad x_3(n) = 2^n \,.$$

The Casoratian for these solutions is:

$$C(n) \;=\; \det\begin{bmatrix} 1 & (-3)^n & 2^n \\ 1 & (-3)^{n+1} & 2^{n+1} \\ 1 & (-3)^{n+2} & 2^{n+2} \end{bmatrix} \;=\; -20 \cdot 2^n \cdot (-3)^n \,,$$

so these solutions are linearly independent.

Theorem 3.7 (Abel)**.** *For $n \geq n_0$ we have*

$$C(n) \;=\; (-1)^{k(n-n_0)}\Big(\prod_{i=n_0}^{n-1} p_k(i)\Big)\, C(n_0)\,, \tag{3.9}$$

where p_k is the coefficient of equation (3.7).

Proof. We can show that $C(n) = (-1)^k p_k(n-1)C(n-1)$. Applying this relation recursively, the above result follows. □

Corollary. *Suppose that $p_k(n) \neq 0$ for all $n \geq n_0$. Then $C(n) \neq 0$ iff $C(n_0) \neq 0$.*

Proof. From the form of (3.9) the statement is immediate. □

From the above corollary we can directly prove the following theorem:

Theorem 3.8. *The set of solutions $x_1(n), x_2(n), \ldots, x_k(n)$ is a fundamental set iff $C(n_0) \neq 0$ for some $n_0 \in \mathbb{N}^*$.*

Example 3.5. The equation

$$a_{n+2} - \frac{3n-2}{n-1} a_{n+1} + \frac{2n}{n-1} a_n = 0$$

has solutions $x_1(n) = n$ and $x_2(n) = 2^n$. We have

$$C(n) \;=\; \det\begin{bmatrix} n & 2^n \\ n+1 & 2^{n+1} \end{bmatrix} .$$

Therefore

$$C(0) \;=\; \det\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \;=\; -1 \neq 0 \,.$$

Thus $\{n, 2^n\}$ is a fundamental set of solutions.

Theorem 3.9. *If $p_k(n) \neq 0$ for all $n \geq n_0$, then the difference equation has a fundamental set for $n \geq n_0$.*

Proof. If $p_k(n) \neq 0$ for all $n \geq n_0$, then $C(n) \neq 0$ for all $n \geq n_0$ and there exist k linearly independent solutions. Thus, there exists a fundamental set of solutions. □

Now we can take advantage of the linearity of linear difference equations to prove the following:

Lemma 3.1. *If $x_1(n), x_2(n)$ are two solutions to a homogeneous linear difference equation, then $x_1(n) + x_2(n)$ and $a\,x_1(n)$ are also solutions, for any constant a.*

Using this result we can show:

Lemma 3.2. *If $x_1(n), x_2(n), \ldots, x_r(n)$ are solutions to a homogeneous linear difference equations then $a_1x_1(n) + a_2x_2(n) + \cdots + a_rx_r(n)$ is also a solution.*

Theorem 3.10. *Let $x_1(n), x_2(n), \ldots, x_k(n)$ be a fundamental set of solutions of the homogeneous equation. Then the general solution for the equation is*

$$x(n) \;=\; \sum_{i=1}^{k} \beta_i\, x_i(n)\,, \tag{3.10}$$

where β_i are constants.

Proof. For any such solution $x(n)$ of the difference equation, we will show that there are β_i's such that (3.10) is true.

If (3.10) is true, then

$$\begin{bmatrix} x_1(n) & x_2(n) & \cdots & x_k(n) \\ x_1(n+1) & x_2(n+1) & \cdots & x_k(n+1) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(n+k-1) & x_2(n+k-1) & \cdots & x_k(n+k-1) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} x(n) \\ x(n+1) \\ \vdots \\ x(n+k-1) \end{bmatrix}.$$

The $k \times k$ matrix on the left hand side is the Casoratian $C(n)$ of the fundamental set of solutions, and thus an invertible matrix. So we can write:

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \;=\; C(n_0)^{-1} \begin{bmatrix} x(n_0) \\ x(n_0+1) \\ \vdots \\ x(n_0+k-1) \end{bmatrix}.$$

The last equation uniquely determines the constants β_i from the initial data. □

Solving Difference Equations

From the previous subsection, we know that the set of fundamental set of solutions to a linear difference equation determines all solutions. We must thus learn how to obtain such a fundamental set of solutions.

Homogeneous Equations with Constant Coefficients

A linear homogeneous equation with constant coefficients is an equation of the form:

$$a_{n+k} + p_1\, a_{n+k-1} + p_2\, a_{n+k-2} + \cdots + p_k\, a_n \;=\; 0\,, \tag{3.11}$$

with p_i known constants. We shall assume that at least $p_k \neq 0$. This guarantees that a_{n+k} is a function of a_n and therefore the equation is of k-th order.

To find solutions, we assume $a_n = \lambda^n$. Then, plugging this into (3.11), we get:

$$\lambda^k + p_1\lambda^{k-1} + p_2\lambda^{k-2} + \cdots + p_k \;=\; 0\,. \tag{3.12}$$

This equation is known as the **characteristic equation**. Solving this k-th order algebraic equation will give us k roots, λ_i, called the **characteristic roots**. Since $p_k \neq 0$, they are all non-zero (which was assumed in deriving (3.12)). There are two cases that must be considered.

Case 1. All roots $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct.

In this case the set $\{\lambda_1^n, \lambda_2^n, \ldots, \lambda_k^n\}$ forms a fundamental set of solutions. To see this we calculate $C(0)$:

$$C(0) \;=\; \det\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_k \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_k^2 \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \cdots & \lambda_k^{k-1} \end{bmatrix}.$$

This determinant (known as the **Vandermonde determinant**) is equal to

$$C(0) = \prod_{1\le i<j\le k} (\lambda_i - \lambda_j)\,,$$

and therefore it is non-zero. The general solution to the difference equation is then

$$x(n) \;=\; \sum_{i=1}^{k} \beta_i\, \lambda_i^n\,. \qquad \square$$

Example 3.6. Find the solution to the initial value problem:

$$a_{n+2} \;=\; a_{n+1} + a_n$$

with the initial conditions $a_1 = 1,\; a_2 = 1$.

Solution. The characteristic equation for this difference equation is:

$$\lambda^2 - \lambda - 1 = 0 ,$$

with solutions[1]

$$\lambda_{\pm} = \frac{1 \pm \sqrt{5}}{2} .$$

The corresponding fundamental set of solutions is thus

$$\left\{ \left(\frac{1+\sqrt{5}}{2}\right)^n , \quad \left(\frac{1-\sqrt{5}}{2}\right)^n \right\} ,$$

giving the general solution:

$$a_n = \beta_1\Big(\frac{1+\sqrt{5}}{2}\Big)^n + \beta_2\Big(\frac{1-\sqrt{5}}{2}\Big)^n .$$

Using the initial conditions we get:

$$\beta_1 = \frac{1}{\sqrt{5}} , \quad \beta_2 = -\frac{1}{\sqrt{5}} .$$

So the solution is

$$a_n = \frac{1}{\sqrt{5}}\Big[\Big(\frac{1+\sqrt{5}}{2}\Big)^n - \Big(\frac{1-\sqrt{5}}{2}\Big)^n\Big] .$$

This is the sequence of the Fibonacci numbers F_n. If we change the initial conditions to $a_1 = 1, a_2 = 3$ we get the sequence of the Lucas numbers L_n:

$$L_n = \Big(\frac{1+\sqrt{5}}{2}\Big)^n + \Big(\frac{1-\sqrt{5}}{2}\Big)^n . \qquad \square$$

Case 2. The roots $\lambda_1, \lambda_2, \dots, \lambda_r$ have multiplicities $m_1, m_2, \dots, m_r$, $r < k$, respectively.

The fundamental set of solutions to (3.11) is the union of the sets

$$\{\lambda_i^n, n\lambda_i^n, n^2\lambda_i^n, \dots, n^{m_i-1}\lambda_i^n\} ,$$

one for each root λ_i. Using the Casoratian, one can show that these are indeed linearly independent. $\square$

Example 3.7. Find the solution to the initial value problem:

$$a_{n+3} - 7a_{n+2} + 16a_{n+1} - 12a_n = 0 ,$$

with the initial conditions $a_0 = 0$, $a_1 = 1$, $a_2 = 1$.

---

[1]The number $\lambda_+ = \frac{1+\sqrt{5}}{2}$ is known as the **golden ratio** and it is usually denoted by ϕ. (Notice that $\lambda_- = -1/\phi$.) Since antiquity, this number has played an important role in science, geometry, architecture, art, music, aesthetics, and philosophy. There are several good books which introduce the reader to this amazing number: [24, 28, 44].

Solution. The characteristic equation for this difference equation is:

$$\lambda^3 - 7\lambda^2 + 16\lambda - 12 \ = \ 0 \ .$$

The fundamental set of solutions corresponding to the above equation is:

$$\{2^n, n2^n, 3^n\}$$

and the general solution is:

$$a_n \ = \ \alpha_0 2^n + \alpha_1 n 2^n + \alpha_2 3^n \ .$$

Using the initial conditions we get:

$$\alpha_0 = 3 \ , \quad \alpha_1 = 2 \ , \quad \alpha_2 \ = \ -3 \ . \qquad \square$$

One may argue that, in the previous discussion, we pulled the solution out of a hat. We thus present a modified proof that is actually constructive. We do so in the case of a 2nd order difference equation. The generalization for a k-th order equation is straightforward but more lengthy.

Given the difference equation

$$a_{n+2} + p_1\, a_{n+1} + p_2\, a_n \ = \ 0 \ , \quad n = 0, 1, 2, \ \ldots \ ,$$

let λ_1, λ_2 be the roots (which may be equal) of the characteristic equation

$$\lambda^2 + p_1\, \lambda + p_2 \ = \ 0 \ .$$

We can then write the difference equation as

$$(\hat{T} - \lambda_1 \hat{I})(\hat{T} - \lambda_2 \hat{I}) a_n \ = \ 0 \ ,$$

where $\hat{T}$ is the translation operator and $\hat{I}$ the unit operator. We now define the new sequence

$$b_n \ = \ (\hat{T} - \lambda_2 \hat{I}) a_n \ , \quad n = 0, 1, 2, \ \ldots \ , \tag{3.13}$$

which allow us to rewrite it furthermore as

$$(\hat{T} - \lambda_1 \hat{I}) b_n \ = \ 0 \ .$$

The last equation is however very simple: It's a the recursion relation of a geometric progression

$$b_{n+1} = \lambda_1\, b_n \ ,$$

with solution

$$b_n \ = \ b_0\, \lambda_1^n \ .$$

Returning to the defining equation (3.13) of b_n, we see that a_n is a solution to

$$a_{n+1} - \lambda_2 a_n = b_0 \lambda_1^n .$$

This equation is solved in Problem 3.4 of page 65. The final solution is

$$a_n = \lambda_2^n a_0 + b_0 \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2} ,$$

if $\lambda_1 \neq \lambda_2$ and

$$a_{n+1} = \lambda_2^n a_0 + b_0 n \lambda_2^{n-1} ,$$

if $\lambda_1 = \lambda_2$. By redefining the constants, we can can write these solutions as

$$a_n = \beta_1 \lambda_1^n + \beta_2 \lambda_2^n ,$$

and

$$a_n = \beta_1 \lambda_2^n + \beta_2 n \lambda_2^n ,$$

respectively.

Linear Non-homogeneous Equations

For a linear non-homogeneous equation of the form:

$$a(n+k) + p_1 a(n+k-1) + \cdots + p_k a(n) = g(n) ,$$

where $g(n) \not\equiv 0$, we have the following theorem:

Theorem 3.11. *The general solution to a linear non-homogeneous equations may be written as*

$$a(n) = x_p(n) + \sum_{i=1}^{k} \beta_i x_i(n) ,$$

where $\{x_1(n), x_2(n), \ldots, x_k(n)\}$ *is the fundamental set of solutions to the homogeneous equation and* $x_p(n)$ *is a particular solution of the homogeneous equation.*

Proof. It is straightforward to verify that this is a solution of the non-homogeneous equation.

Inversely, let $x_p(n)$ be a solution of the non-homogeneous equation. Then

$$x_p(n+k) + p_1 x_p(n+k-1) + \cdots + p_k x_p(n) = g(n) .$$

Subtracting this from the given difference equation, we see that $a(n) - x_p(n)$ is a solution of the homogeneous equation and therefore

$$a(n) - x_p(n) = \sum_{i=1}^{k} \beta_i x_i(n) . \qquad \square$$

Therefore, the solution of non-homogeneous equations is essentially making an educated guess as to what a solution x_p may be. There are several methods that have been developed to find such a solution. The most useful seems to be **method of undetermined coefficients**. This method works well when $g(n)$ is of the form a^n, $\sin(bn)$, $\cos(bn)$ or n^ℓ (or some combination of them). Table 3.1 gives the form of x_p for each case of $g(n)$. Once a guess has been made, it must be plugged in and the respective coefficients must be solved for.

Forcing term $g(n)$	**Particular solution** x_p
a^n	$c\,a^n$
n^ℓ	$c_0 + c_1 n + \cdots + c_\ell n^\ell$
$a^n n^\ell$	$a^n\,(c_0 + c_1 n + \cdots + c_\ell n^\ell)$
$\sin(bn)$, $\cos(bn)$	$c_1 \sin(bn) + c_2 \cos(bn)$
$a^n \sin(bn)$, $a^n \cos(bn)$	$a^n\ [c_1 \sin(bn) + c_2 \cos(bn)]$
$a^n n^\ell \sin(bn)$, $a^n n^\ell \cos(bn)$	$a^n \left[(c_0 + c_1 n + \cdots + c_\ell n^\ell) \sin(bn) + (d_0 + d_1 n + \cdots + d_\ell n^\ell)\ \cos(bn)\right]$

Table 3.1: Most common forcing terms with the corresponding particular solutions.

Example 3.8. Find the general solution to the following difference equation:

$$a_{n+2} + a_{n+1} - 12a_n \ = \ n\,2^n\ .$$

Solution. Using the technique of the previous section, the solution to the homogeneous equation is

$$\beta_1\,3^n + \beta_2\,(-4)^n\ .$$

Because this is in the form $b^n n^\ell$, we guess that $x_p(n)$ has the form $c_1 2^n + c_2 n 2^n$. Plugging this into the equation and solving for the constants, we get

$$x_p(n) \ = \ -\frac{5}{18}\,2^n - \frac{1}{6}\,n2^n\ .$$

Thus the general solution is

$$a_n \ = \ \beta_1 3^n + \beta_2(-4)^n - \frac{5}{18}\,2^n - \frac{1}{6}\,n2^n\ .$$ □

3.4 Solved Problems

Problem 3.1 (UK 1996; Estonia 1997; Greece 1999)**.** *A function defined on the positive integers satisfies*

(a) $f(1) = 1996$
(b) $f(1) + \cdots + f(n) = n^2 f(n)$, $n > 1$.
Find $f(1996)$.

Solution. In condition (b), substituting $n-1$ in place of n we find

$$f(1) + \cdots + f(n-1) \;=\; (n-1)^2 f(n-1)\,,$$

Then subtracting this equation from the given one, we find a first order difference equation

$$f(n) \;=\; n^2 f(n) - (n-1)^2 f(n-1)\,,$$

which implies

$$f(n) \;=\; \frac{n-1}{n+1}\, f(n-1)\,.$$

By iteration

$$f(n) \;=\; \frac{n-1}{n+1}\,\frac{n-2}{n}\,\frac{n-3}{n-1}\cdots\frac{1}{3}\, f(1) \;=\; \frac{2}{n(n+1)}\, f(1)\,.$$

Therefore

$$f(1996) \;=\; \frac{2}{1996\cdot 1997}\,1996 \;=\; \frac{2}{1997}\,. \qquad \square$$

Problem 3.2. *The function F is defined in the set of natural numbers* $\mathbb{N}^*$ *and satisfies the conditions*
(a) $F(n) = F(n-1) + a^n$,
(b) $F(1) = 1$.
Find $F(n)$.

Solution. In condition (a), in place of n we substitute $2, 3, \ldots, n-1, n$:

$$\begin{aligned} F(2) &= F(1) + a^2 \\ F(3) &= F(2) + a^3 \\ &\cdots \\ F(n-1) &= F(n-2) + a^{n-1} \\ F(n) &= F(n-1) + a^n \end{aligned}$$

Adding these equation we find

$$F(n) \;=\; F(1) + \sum_{k=2}^{n} a^k\,.$$

If $a \neq 1$, then

$$F(n) \;=\; 1 + \frac{a^2(a^{n-1}-1)}{a-1}\,.$$

If $a = 1$, then

$$F(n) \;=\; n\,. \qquad \square$$

Problem 3.3. *A sequence of real numbers a_n satisfying the condition*

$$a_n = \lambda\, a_{n-1} + \omega\,, \quad n = 0, 1, 2, \ldots\,,$$

with $\lambda \neq 0, 1$ and $\omega \neq 0$, is given. Find the expression of a_n in terms of a_0.

Comment. This sequence is known as a **mixed progression** with ratio λ and step ω.

Solution. In the given condition, in place of n we substitute $1, 2, 3, \ldots, n-1, n$:

$$\begin{aligned} a_1 &= \lambda\, a_0 + \omega \\ a_2 &= \lambda\, a_1 + \omega \\ a_3 &= \lambda\, a_2 + \omega \\ &\cdots \\ a_{n-1} &= \lambda\, a_{n-2} + \omega \\ a_n &= \lambda\, a_{n-1} + \omega \end{aligned}$$

We multiply them by $\lambda^{n-1}, \lambda^{n-2}, \ldots, \lambda, 1$ respectively and then add the resulting equations to find:

$$a_n \;=\; \lambda^n\, a_0 + \omega \sum_{k=0}^{n-1} \lambda^k\,,$$

or

$$a_n \;=\; \lambda^n\, a_0 + \omega\, \frac{\lambda^n - 1}{\lambda - 1}\,. \qquad \square$$

Problem 3.4. *The sequence of real numbers a_n satisfying the condition*

$$a_n = \lambda\, a_{n-1} + \omega\, \mu^{n-1}\,, \quad n = 0, 1, 2, \ldots\,,$$

with $\lambda \neq 0$, generalizes the mixed progression of the previous problem. Find the expression of a_n in terms of a_0.

Solution. In the given condition, in place of n we substitute $1, 2, 3, \ldots, n-1, n$:

$$\begin{aligned} a_1 &= \lambda\, a_0 + \omega \\ a_2 &= \lambda\, a_1 + \omega\, \mu \\ a_3 &= \lambda\, a_2 + \omega\, \mu^2 \\ &\cdots \\ a_{n-1} &= \lambda\, a_{n-2} + \omega\, \mu^{n-2} \\ a_n &= \lambda\, a_{n-1} + \omega\, \mu^{n-1} \end{aligned}$$

We multiply them by λ^{n-1}, λ^{n-2}, λ^{n-1}, λ, 1 respectively and then we add the resulting equations to find:

$$a_n \;=\; \lambda^n\, a_0 + \frac{\omega}{\lambda}\mu^n \sum_{k=1}^{n}\left(\frac{\lambda}{\mu}\right)^k \;.$$

If $\mu = \lambda$,

$$a_n \;=\; \lambda^n\, a_0 + \omega\, n\lambda^{n-1} \;,$$

whereas if $\mu \neq \lambda$,

$$a_n \;=\; \lambda^n\, a_0 + \omega\,\frac{\lambda^n - \mu^n}{\lambda - \mu} \;. \qquad \square$$

Problem 3.5. *A sequence of real numbers a_n satisfying the condition*

$$a_n = \lambda\, x^{a_{n-1}} + \omega \;,$$

where $x > 0$ and $\lambda \neq 0$, is given. Find the expression of a_n in terms of a_0.

Solution. The given recursion relation may be written in the form

$$a_n - \omega \;=\; \lambda x^{\omega}\, x^{a_{n-1}-\omega} \;.$$

To simplify the notation we introduce

$$\beta_n \;=\; a_n - \omega \;, \quad \mu \;=\; \lambda x^{\omega} \;.$$

Then

$$\beta_n \;=\; \mu\, x^{\beta_{n-1}} \;.$$

If $\mu = 1$ by iteration we can easily find

$$\beta_n \;=\; x^{x^{\cdot^{\cdot^{\cdot^{x^{\beta_0}}}}}}$$

where the x appears n times. If $\mu \neq 1$, then we write

$$\frac{\beta_n}{\mu} \;=\; (x^{\mu})^{\frac{\beta_{n-1}}{\mu}} \;,$$

and we define

$$\gamma_n \;=\; \frac{\beta_n}{\mu} \;, \quad y \;=\; x^{\mu} \;.$$

Therefore

$$\gamma_n \;=\; y^{\gamma_{n-1}}$$

with solution

$$\gamma_n \;=\; y^{y^{\cdot^{\cdot^{\cdot^{y^{\gamma_0}}}}}} \;,$$

where the y appears n times. This can be rewritten for the original sequence a_n:

$$a_n \;=\; \omega + \lambda x^{\omega}\, y^{y^{\cdot^{\cdot^{\cdot^{y^{\frac{a_0-\omega}{\lambda x^{\omega}}}}}}}} \;, \quad y = x^{\lambda x^{\omega}} .$$

□

Problem 3.6 (IMO 1976)**.** *A sequence (u_n) is defined by*

$$\begin{aligned} u_0 &= 2\,, \quad u_1 \;=\; 5/2\,, \\ u_{n+1} &= u_n\,(u_{n-1}^2 - 2) - u_1\,, \qquad n \geq 1, 2, \cdots \,. \end{aligned}$$

Prove that for positive integers n,

$$\lfloor u_n \rfloor \;=\; 2^{\frac{\lfloor 2^n-(-1)^n \rfloor}{3}} \;,$$

where $\lfloor x \rfloor$ denotes the integer part of x.

Solution. Setting $n = 1, 2$ in the given recursion relation we find $u_2 = 5/2$ and $u_3 = 65/8$. Then we notice that

$$\begin{aligned} u_0 &= 2 \;=\; 2^0 + 2^{-0}\,, \\ u_1 &= \frac{5}{2} \;=\; 2^1 + 2^{-1}\,, \\ u_2 &= \frac{5}{2} \;=\; 2^1 + 2^{-1}\,, \\ u_3 &= \frac{65}{8} \;=\; 2^3 + 2^{-3}\,. \end{aligned}$$

We thus suspect that u_n can be written in the form

$$u_n \;=\; 2^{a_n} + 2^{-a_n} \;,$$

for some sequence $\{a_n\}$. Even more, it appears that the first term will be an integer and the second term fractional, implying that the integer part of u_n is solely based on its first term. We thus come to believe that[2]

$$a_n \;=\; \frac{2^n - (-1)^n}{3} \;,$$

which will prove by the method of induction.

[2]Notice that 3 always divides $2^n - (-1)^n$. You may try to show this formally if it is not evident or known to you.

The statement is true for the first terms as we have seen. Let it be true for all terms up to some index k. Then if

$$u_{k+1} = 2^{a_{k+1}} + 2^{-a_{k+1}}, \quad a_{k+1} = \frac{2^{k+1} - (-1)^{k+1}}{3},$$

we will show that u_{k+1} satisfies the recursion relation of the problem. First we see that

$$\begin{aligned} u_k\,(u_{k-1}^2 - 2) &= (2^{a_k} + 2^{-a_k})\left[(2^{a_{k-1}} + 2^{-a_{k-1}})^2 - 2\right] \\ &= (2^{a_k} + 2^{-a_k})\left(2^{2a_{k-1}} + 2^{-2a_{k-1}}\right) \\ &= 2^{a_k+2a_{k-1}} + 2^{-(a_k+2a_{k-1})} + 2^{a_k-2a_{k-1}} + 2^{-(a_k-2a_{k-1})}. \end{aligned}$$

Then

$$\begin{aligned} a_k + 2a_{k-1} &= \frac{2^k - (-1)^k}{3} + 2\,\frac{2^{k-1} - (-1)^{k-1}}{3} \\ &= \frac{2^k + (-1)^{k+1}}{3} + \frac{2^k - 2\,(-1)^{k+1}}{3} = \frac{2^{k+1} - (-1)^{k+1}}{3} = a_{k+1}, \end{aligned}$$

and

$$\begin{aligned} a_k - 2a_{k-1} &= \frac{2^k - (-1)^k}{3} - 2\,\frac{2^{k-1} - (-1)^{k-1}}{3} \\ &= \frac{2^k + (-1)^{k+1}}{3} - \frac{2^k - 2\,(-1)^{k+1}}{3} = (-1)^{k+1}. \end{aligned}$$

Therefore

$$u_k\,(u_{k-1}^2 - 2) = 2^{a_{k+1}} + 2^{-a_{k+1}} + 2^{(-1)^{k+1}} + 2^{-(-1)^{k+1}}.$$

Notice that the last two terms are 2^1 or 2^{-1} for all k. And their sum is u_1. We thus conclude that

$$u_k\,(u_{k-1}^2 - 2) = u_{k+1} + u_1,$$

and this concludes the proof. □

Chapter 4

Equations Reducing to Algebraic Systems

In this chapter we study functional equations and systems of functional equations that can be solved by a reduction to a set of algebraic equations.

4.1 Solved Problems

Problem 4.1. *Find an even function $f : \mathbb{R} \to \mathbb{R}$ and an odd function $g : \mathbb{R} \to \mathbb{R}$ such that $2^x = f(x) + g(x), \forall x \in \mathbb{R}$.*

Solution. Substituting x by $-x$ in the given equation, we get $2^{-x} = f(x) - g(x)$. Solving the system of this equation and the original one, we obtain

$$\begin{aligned} f(x) &= \frac{2^x + 2^{-x}}{2} = \cosh(x \ln 2)\,, \\ g(x) &= \frac{2^x - 2^{-x}}{2} = \sinh(x \ln 2)\,. \end{aligned}$$

□

Problem 4.2. *Find $f : \mathbb{R}^* \to \mathbb{R}$ such that*

$$f(x) + 2 f\left(\frac{1}{x}\right) = x\,, \quad \forall x \in \mathbb{R}^*\,.$$

Solution. In the defining equation we set $y = 1/x$:

$$f\left(\frac{1}{y}\right) + 2 f(y) = \frac{1}{y}\,.$$

This and the given equation give the system of equations

$$\begin{aligned} f(x) + 2f\left(\frac{1}{x}\right) &= x\,, \\ f\left(\frac{1}{x}\right) + 2f(x) &= \frac{1}{x}\,. \end{aligned}$$

This can be solved easily to give

$$f(x) = \frac{2-x^2}{3x}\,. \qquad \square$$

Problem 4.3. *Find* $f : \mathbb{R} \setminus \{0,1\} \to \mathbb{R}$ *such that*

$$f\left(\frac{x}{x-1}\right) - 2f\left(\frac{x-1}{x}\right) = 0\,, \quad \forall x \in \mathbb{R} \setminus \{0,1\}\,.$$

Solution. In the defining equation we set $y = x/(x-1)$. Then

$$f(y) - 2f\left(\frac{1}{y}\right) = 0\,.$$

Now we set $t = 1/y$:

$$f\left(\frac{1}{t}\right) - 2f(t) = 0\,,$$

The last two equations give

$$f(t) \equiv 0\,. \qquad \square$$

Problem 4.4. *Find* $f : \mathbb{R} \setminus \{0,1\} \to \mathbb{R}$ *such that*

$$f(x) + f\left(\frac{1}{1-x}\right) = x\,, \quad \forall x \in \mathbb{R} \setminus \{0,1\}\,. \tag{4.1}$$

Solution. In the defining equation (4.1) we set $y = 1/(1-x) \Leftrightarrow x = (y-1)/y$. Then

$$f\left(\frac{y-1}{y}\right) + f(y) = \frac{y-1}{y}\,.$$

From this equation and the given one (4.1), we find

$$f\left(\frac{1}{1-x}\right) - f\left(\frac{x-1}{x}\right) = x - \frac{x-1}{x}\,. \tag{4.2}$$

In (4.2) we set again $y = 1/(1-x) \Leftrightarrow x = (y-1)/y$:

$$f(y) - f\left(\frac{1}{1-y}\right) \;=\; \frac{y-1}{y} - \frac{1}{y-1}\;. \tag{4.3}$$

Equations (4.1) and (4.3) can be solved for $f(x)$ easily to give

$$f(x) \;=\; \frac{x^3 - x + 1}{2x^2 - 2x}\;. \qquad \square$$

Problem 4.5 (Putnam 1959)**.** *Find all* $f : \mathbb{C} \to \mathbb{C}$ *such that*

$$f(z) + z\,f(1-z) \;=\; 1 + z\,, \quad \forall z \in \mathbb{C}\,. \tag{4.4}$$

Solution. In the defining equation (4.4) we set $w = 1 - z \Leftrightarrow z = 1 - w$. Then

$$f(1-w) + (1-w)\,f(w) \;=\; 2 - w\,. \tag{4.5}$$

From (4.4) and (4.5) we find easily

$$(1 - z + z^2)\,f(z) = (1 - z + z^2)\,. \tag{4.6}$$

Therefore if

$$1 - z + z^2 \neq 0 \;\Leftrightarrow\; z \neq \frac{1 \pm i\sqrt{3}}{2} \;\equiv\; \rho_\pm\,,$$

equation (4.6) implies that

$$f(z) \;=\; 1\,, \quad \forall z \in \mathbb{C} \setminus \{\rho_+, \rho_-\}\,.$$

For $z = \rho_\pm$, equation (4.6) is an identity. However, the defining relation (4.4) gives

$$\begin{aligned} f(\rho_+) + \rho_+\,f(1-\rho_+) \;&=\; 1 + \rho_+\,, \\ f(\rho_-) + \rho_-\,f(1-\rho_-) \;&=\; 1 + \rho_-\,. \end{aligned}$$

Only one of these two equations is independent. This can be seen by observing $\rho_+ + \rho_- = 1$, $\rho_+\rho_- = 1$ and multiplying the first, for example, by ρ_-; then we recover the second. So, the value of f at one of the $\rho_\pm$ is arbitrary. Let's say

$$f(\rho_-) \;=\; z_0\,.$$

Then we find

$$f(\rho_+) \;=\; 1 + \rho_+\,(1 - z_0)\,. \qquad \square$$

Problem 4.6. *Let* $f : \mathbb{R} \to \mathbb{R}$ *such that*

$$f(x + y) + f(x - y) \ = \ 2\, f(x)\, \cos y\,, \quad \forall x, y \in \mathbb{R}\,. \tag{4.7}$$

Find all functions f.

Solution. Setting $x = 0$ and $y = \theta$ in the defining relation we find

$$f(\theta) + f(-\theta) \ = \ 2a\, \cos\theta\,,$$

where $a = f(0)$.

Setting $x = \theta + \pi/2$ and $y = \pi/2$ in the defining relation we find

$$f(\theta + \pi) + f(\theta) \ = \ 0\,.$$

Finally, setting $x = \pi/2$ and $y = \theta + \pi/2$ in the defining relation we find

$$f(\theta + \pi) + f(-\theta) \ = \ -2b\, \sin\theta\,,$$

where $b = f(\pi/2)$.

We have thus derived three equations for the three unknowns $f(\theta)$, $f(-\theta)$, $f(\theta + \pi)$. The solution is easily found to be

$$f(\theta) \ = \ a\, \cos\theta + b\, \sin\theta\,. \qquad \square$$

Question. *Let* $f : \mathbb{R} \to \mathbb{R}$ *such that*

$$f(x + y) + f(x - y) \ = \ 2\, f(x)\, \cosh y\,, \quad \forall x, y \in \mathbb{R}\,. \tag{4.8}$$

Why doesn't the above method work for finding f*?*

Answer. With the substitution pointed out in the solution of equation (4.7), we were able to find a system of three equations for three unknowns $f(\theta)$, $f(-\theta)$, $f(\theta + \pi)$. In particular, the substitution $y = \pi/2$ in (4.7) makes the right hand side to vanish. This side would contain $f(\theta + \pi/2)$, a fourth unknown, otherwise. However the modified equation contains the hyperbolic cosine that does not vanish for any finite real value of its argument.

Solution of Equation (4.8). To solve equation (4.8) one can extend the domain of the definition of f from $\mathbb{R}$ to $\mathbb{C}$ by *analytic continuation* and define $f : \mathbb{C} \to \mathbb{C}$ such that

$$f(z + w) + f(z - w) \ = \ 2\, f(z)\, \cosh w\,, \quad \forall z, w \in \mathbb{C}\,.$$

Setting $(z, w) = (0, it)$, $(z, w) = (i(t + \pi/2), i\pi/2)$, $(z, w) = (i\pi/2, i(t + \pi/2))$, successively, we find:

$$\begin{aligned} f(it) + f(-it) \ &= \ 2a\, \cos t\,, \\ f(i(t + \pi)) + f(-it) \ &= \ 2a\, \cos t\,, \\ f(i(t + \pi)) + f(-it) \ &= \ -2b\, \sin t\,, \end{aligned}$$

where $a = f(0)$, $b = f(i\pi/2)$.
We have thus, again, derived three equations for the three unknowns $f(it)$, $f(-it)$, $f(i(t+\pi))$. The solution is easily found

$$f(it) \;=\; a\, \cos t + b\, \sin t\, .$$

Returning to real values $\theta = it$:

$$f(\theta) \;=\; a\, \cosh\theta + ib\, \sinh\theta\, .$$

f will be real-valued if a is real and b is imaginary. □

Problem 4.7 ([1], Problem 11030)**.** *Show that for $d < -1$ there are exactly two real-valued functions f such that*

$$f(x+y) - f(x)f(y) \;=\; d\, \sin x\, \sin y\, , \quad \forall x, y \in \mathbb{R}\, .$$

Solution. Letting $x = 0 = y$ we get $f(0)(1 - f(0)) = 0$. Therefore, $f(0)$ can be 0 or 1. If $f(0) = 0$, then by setting $y = 0$ we get $f(x) = 0, \forall x \in \mathbb{R}$. But $f(x) \equiv 0$ is clearly not a solution, thus $f(0) = 1$.

Replacing (x, y) by $(t, t), (t, -t), (2t, -t)$ sequentially, we obtain the system

$$\begin{aligned} f(2t) &= (f(t))^2 + d\sin^2 t\, , \\ 1 &= f(t)f(-t) - d\sin^2 t\, , \\ f(t) &= f(2t)f(-t) - 2d\sin^2 t\cos t\, . \end{aligned}$$

Eliminating $f(-t)$ and $f(2t)$ we get

$$d\, \sin^2 t\, [(f(t))^2 - 2\cos t f(t) + (1 + d\sin^2 t)] \;=\; 0\, .$$

If $t \neq n\pi$, with n being an integer, then $f(t) = \cos t \pm \sqrt{-d-1}\sin t$ by the quadratic formula.

It remains only to find the values $f(n\pi)$. Setting $t = \pi/2$ in the first equation of the system we get

$$f(\pi) = (f(\frac{\pi}{2}))^2 + d \;=\; -d - 1 + d = -1\, .$$

Setting $t = \pi$ in the second equation of the system, we get $f(-\pi) = -1$. Finally, inductively from the given functional equation, we can easily obtain that $f(2k\pi) = 1$ and $f((2k+1)\pi) = -1$ for $k \in \mathbb{Z}$. □

Problem 4.8. *Find* $f, g : \mathbb{R} \setminus \{1\} \to \mathbb{R}$ *such that*

$$\begin{aligned} f(2x+1) + 2\,g(2x+1) &= 2x\,, &(4.9)\\ f\left(\frac{x}{x-1}\right) + g\left(\frac{x}{x-1}\right) &= x\,. &(4.10) \end{aligned}$$

Solution. In the equation (4.9) we set $t = 2x + 1$:

$$f(t) + 2\,g(t) = t - 1\,.$$

In the equation (4.10) we set $t = \frac{x}{x-1}$:

$$f(t) + g(t) = \frac{t}{t-1}\,.$$

The last two equations give

$$f(x) = \frac{x^2 - 4x + 1}{-x + 1}\,, \quad g(x) = \frac{x^2 - 3x + 1}{x - 1}\,.$$ □

4.2 Group Theory in Functional Equations

Group theory is an essential tool for many branches of mathematics. In this section we shall present the basic ideas and then how they can be used in the solution of some functional equations. To motivate what will be presented in this section, let's solve again the Problem 4.4.

Problem 4.9. *Find* $f : \mathbb{R} \setminus \{0, 1\} \to \mathbb{R}$ *such that*

$$f(x) + f\left(\frac{1}{1-x}\right) = x\,, \quad \forall x \in \mathbb{R} \setminus \{0, 1\}\,.$$

Solution. Note that in this problem, if we let $g(x) = 1/(1-x)$, then $g^2(x) = (g \circ g)(x) = (x-1)/x$, and $g^3(x) = x$. The defining equation (4.1), together with the substitutions of x by $g(x)$ and $g^2(x)$, yield the system

$$\begin{aligned} f(x) + f(g(x)) &= x\,,\\ f(g(x)) + f(g^2(x)) &= g(x)\,,\\ f(g^2(x)) + f(x) &= g^2(x)\,. \end{aligned}$$

Solving the system, we get

$$f(x) = \frac{x - g(x) + g^2(x)}{2} = \frac{x^3 - x + 1}{2x^2 - 2x}\,.$$ □

So what's important about this solution? Consider the set $G = \{g, g^2, g^3\}$. When this set is equipped with the operation of function composition, then some interesting properties emerge. In particular, composition of any three functions satisfies the associative property $f \circ (h \circ k) = (f \circ h) \circ k$; the $g^3(x)$ element is an identity element — that is, when composed with any function f, it does not change the function $g^3 \circ f = f \circ g^3 = f$; and, for any function f, there is another function f' such that their composition is the identity function, $f \circ f' = f' \circ f = g^3$. In particular, for $f = g$, $f' = g^2$, for $f = g^2$, $f' = g$, and for $f = g^3$, $f' = g^3$. To describe the structure that has been revealed in G , we say that G equipped with the operation of composition is a *group*.

All this should strike a familiar cord. Recall the set of integers $\{..., -2, -1, 0, 1, 2, ...\}$. As a set it is not a very interesting object since there is no relation between its elements. However, when this set is equipped with the operation of addition, then interesting properties (as above) emerge. In particular, addition of any three integers satisfies the associative property $m + (n + \ell) = (m + n) + \ell$; the 0 element, when added to any integer n, does not change the integer: $0 + n = n + 0 = n$; and, for any integer n, there is another integer $-n$ such that their sum is 0: $n + (-n) = (-n) + n = 0$. $\mathbb{Z}$ equipped with the operation of addition is a group. Moreover, for any two integers m and n, $m + n = n + m$. We say that the operation of addition is commutative and that $\mathbb{Z}$ is a *commutative group*.

In the following part of the section we present some general ideas about groups and we use them in some functional equations. Simple group theoretic ideas will also be met in some of the later chapters.

Elementary Group Theory

Group theory is an important subject in mathematics due to its many implications and applications. Some of its most important applications are in physics where it is used to describe symmetries. In this section we show how one may use group theory as a tool to aid us in finding solutions to certain functional equations.

Definition 4.1. Given a set G and a binary operation $*$ over the set G that is closed (meaning that if $x \in G$ and $y \in G$ then $x * y \in G$) and satisfies the following axioms

1. *Associativity*: $x * (y * z) = (x * y) * z$ for all $x, y, z \in G$;
2. *Neutral element*: there exists an $e \in G$ such that $e * x = x * e = x$ for all $x \in G$;
3. *Symmetric elements*: for each $x \in G$, there exists a $x' \in G$ such that $x' * x = x * x' = e$;

we say that G endowed with $*$ is a **group**. If $*$ satisfies in addition the property

4. *Commutativity*: $x * y = y * x$, for $x, y \in G$,

then the group is called an **abelian** or **commutative** group.

For an additive operation we call the neutral element *the zero* and the symmetric elements *the opposites*. For a multiplicative operation we call the neutral element *the unit* and the symmetric elements *the inverses*.

As we have seen, the set $\mathbb{Z}$ endowed with the ordinary addition is an abelian group. Some additional examples are given below.

Example 4.1. (i) ***The single element group***: The group consisting of the identity element alone is a group. Since we have $G = \{e\}$ and $e \circ e = e$ it is closed under $\circ$ and e is its own inverse, while associativity and the neutral element axioms are satisfied trivially. Thus G is a group.

(ii) ***Rational numbers***: The set of integers together with multiplication cannot be a group since the inverse of $a \in \mathbb{Z}$ is $1/a$ which is not an integer. However all the set of all rational numbers, not including zero, of the form a/b where $a, b \in \mathbb{Z}$ with multiplication do form a group. It is clear that closure is satisfied since $\frac{a}{b} \circ \frac{c}{d} = \frac{ac}{bd}$ for $a, b, c, d \in \mathbb{Z}$. Since multiplication is associative for integers it must still be associative for rationals. The identity axiom is clearly satisfied by 1. Lastly, the inverse of $a/b \in \mathbb{Q}$ is obviously b/a. We can also see that since multiplication is commutative for integers it is also commutative for the rationals. Thus the rationals together with multiplication form an abelian group.

(iii) ***Non-singular matrices***: It can be shown from matrix theory that if A and B are non-singular matrices, i.e. $\det(A) \neq 0$, then the matrix product AB is also non-singular. The matrix product is defined to be associative and the identity matrix I has determinant 1, so it is non-singular. Finally, since the requirements for a matrix to be invertible is that it be non-singular, all nonsingular matrices have an inverse. Thus the group of non-singular matrices under matrix multiplication form a group.

(iv) ***Cyclic groups***: A special type of group that is of interest to functional equations is the cyclic groups. Cyclic groups are groups such that each element can be made from repeated products of the same element, i.e. there exists an element, $h \in G$ such that each element can be expressed by h^i where $i \in \mathbb{N}$. We can see that if there are n elements in the group, then h^n must be the identity. So the set G is $G = \{h, h^2, h^3, \ldots, h^{n-1}, h^n = e\}$.

Definition 4.2. We say that the set $A = \{a_1\, a_2, \ldots, a_n\}$ **generates** the group G with operation $*$ if every element g of G can be written in the form $g = a_1^{m_1} a_2^{m_2} \ldots a_n^{m_n}$ where $m_1, m_2, \ldots, m_n$ are integers.

Definition 4.3. The **order** of an element $g \in G$, if it exists, is the smallest natural number n such that $g^n = e$. The **order** of a group G, if it is finite, is the number of the elements in G.

Each element of a finite group G has a finite order. If G is a cyclic group with n elements, then its order and every element's order is n.

Definition 4.4. A **subgroup** H of a group $(G, *)$ is a subset of G which is also a group under the operation $*$.

Every group G has at least two trivial subgroups: the entire group G and the single element group $\{e\}$.

Theorem 4.1 (Lagrange)**.** *For a finite group G, the order of a subgroup H divides the order of the group.*

The ratio of the two orders we call **index** of H in G and denote by $[G : H]$.

Group tables are a way to represent a finite group and its structure. The rows and columns are labelled by an element in the group and the (i, j)-th entry of the table is the element made by operating the element in the i-th row by the element in the j-th column. A simple example is the group $G = \{-1, 1\}$ under ordinary multiplication:

·	-1	1
-1	1	-1
1	-1	1

Using group tables is an easy way to recognize properties of the group. For instance, if a group is abelian, and the rows and columns are ordered in the same way then for each pair i, j the (i, j)-th element will be equal to the (j, i)-th element (i.e. the table is symmetric about the diagonal).

Before we finish this section, let's look at the Problem 1.2 again and make some connections with group theory. The set G defined there, equipped with the operation of composition of functions, is easily seen to be a group. The element $f \circ g \circ f^{-1} \circ g^{-1}$ that was used in the solution of the problem is called the **commutator** of f and g and it is often denoted by $[f, g]$. If a group is abelian, then all elements commute and all commutators reduce to the neutral element. Therefore, commutators measure the deviation of a group from being abelian. We can make this concept more precise as follows: We take all possible commutators of a group and form a set. Unfortunately, in general, this set does not consist a subgroup of G. However, we can consider it as a set that generates a subgroup denoted by $[G, G]$ and called the **commutator subgroup** or **derived subgroup**. The following theorem is then true:

Theorem 4.2. *A group is abelian iff its commutator subgroup is the single element group $\{e\}$.*

Condition (c) of Problem 1.2 was requiring that the group G is abelian. This is obvious from the commutator $[f, g]$ which is the identity function or it can be seen explicitly by a simple calculation:

$$\begin{aligned} f \circ g(x) &= A\,x + B\,, \\ g \circ f(x) &= A\,x + C\,, \end{aligned}$$

where $A = aa'$, $B = ab' + b$, $C = a'b + b'$. Condition (c) imposes $B = C$. (Recall the solution of Problem 1.2.)

Using Group Theory to Reduce Functional Equations to Algebraic Systems

If we have a set $G = \{g_1, ..., g_n\}$ of mappings from a subset of $\mathbb{C}$ to $\mathbb{C}$ which becomes a group when it is endowed with the operation of map composition, then we can ask what are the solutions $f(z)$ of the functional equation

$$a_1(z)\, f(g_1(z)) + a_2(z)\, f(g_2(z)) + ... + a_n(z)\, f(g_n(z)) \;=\; b(z)\,,$$

where $a_1(z), \ldots, a_n(z)$, and $b(z)$ are some given functions.

In such a case the solution to the functional equation can be found easily as follows. Substituting z in the defining equation by $g_1(z), \ldots, g_n(z)$, we obtain a system of n equations with the n unknowns being $f(g_1(z)), \ldots, f(g_n(z))$. We can then solve the system for $f(z)$ at all the values of z where the coefficient matrix is non-singular. At the points where the coefficient matrix is singular additional work is necessary. This latter comment is better understood by looking again at the Putnam 1959 problem.

To avoid cumbersome notation, let's take G to be a cyclic group of order n. Then $G = \{e, g, g^2, \ldots, g^{n-1}\}$ with $g^n = e$. Then the algebraic system found after the substitution of z by $g(z), g^2(z), \ldots, g^{n-1}(z)$ is

$$\begin{aligned}
a_1(g(z))\, f(g(z)) + a_2(g(z))\, f(g^2(z)) + \ldots + a_n(g(z))\, f(z) &= b(g(z))\,, \\
a_1(g^2(z))\, f(g^2(z)) + a_2(g^2(z))\, f(g^3(z)) + \ldots + a_n(g^2(z))\, f(g(z)) &= b(g^2(z))\,, \\
\cdots \quad &\cdots \quad \cdots\,, \\
a_1(g^{n-1}(z))\, f(g^{n-1}(z)) + a_2(g^{n-1}(z))\, f(z) + \ldots + a_n(g^{n-1}(z))\, f(g^{n-2}(z)) &= b(g^{n-1}(z))\,, \\
a_1(z)\, f(z) + a_2(z)\, f(g(z)) + \ldots + a_n(z)\, f(g^{n-1}(z)) &= b(z)\,,
\end{aligned}$$

or, after rearranging the terms,

$$\begin{aligned}
a_n(g(z))\, f(z) + a_1(g(z))\, f(g(z)) + \ldots + a_{n-1}(g(z))\, f(g^{n-1}(z)) &= b(g(z))\,, \\
a_{n-1}(g^2(z))\, f(z) + a_n(g^2(z))\, f(g(z)) + \ldots + a_{n-2}(g^2(z))\, f(g^{n-1}(z)) &= b(g^2(z))\,, \\
\cdots \quad &\cdots \quad \cdots\,, \\
a_2(g^{n-1}(z))\, f(z) + a_3(g^{n-1}(z))\, f(g(z)) + \ldots + a_1(g^{n-1}(z))\, f(g^{n-1}(z)) &= b(g^{n-1}(z))\,, \\
a_1(z)\, f(z) + a_2(z)\, f(g(z)) + \ldots + a_n(z)\, f(g^{n-1}(z)) &= b(z)\,.
\end{aligned}$$

We can then solve the system for $f(z)$ at all the values of z where the matrix

$$A(z) = \begin{bmatrix} a_n(g(z)) & a_1(g(z)) & \ldots & a_{n-1}(g(z)) \\ a_{n-1}(g^2(z)) & a_n(g^2(z) & \ldots & a_{n-2}(g^2(z)) \\ \vdots & \vdots & \vdots & \vdots \\ a_2(g^{n-1}(z)) & a_3(g^{n-1}(z)) & \ldots & a_n(g^{n-1}(z)) \\ a_1(g^n(z)) & a_2(g^n(z)) & \ldots & a_n(g^n(z)) \end{bmatrix}$$

satisfies $\det A \neq 0$. Then

$$f(z) = \frac{\det B(z)}{\det A(z)}\,,$$

where

$$B(z) = \begin{bmatrix} b(g(z)) & a_1(g(z)) & \ldots & a_{n-1}(g(z)) \\ b(g^2(z)) & a_n(g^2(z) & \ldots & a_{n-2}(g^2(z)) \\ \vdots & \vdots & \vdots & \vdots \\ b(g^{n-1}(z)) & a_3(g^{n-1}(z)) & \ldots & a_n(g^{n-1}(z)) \\ b(g^n(z)) & a_2(g^n(z)) & \ldots & a_n(g^n(z)) \end{bmatrix}.$$

In general, the algebraic and computational work involved in solving such systems can be very demanding for larger n and non-constant coefficients.

Having understood the ideas presented in this section, the reader may want to return to some of the functional equations of the previous section and rethink them in the new language. Below, we illustrate the method with an additional example in which the underlying group has order 6 but is not cyclic.

Problem 4.10. *Find all functions* $f : \mathbb{C} \setminus \{0,1\} \to \mathbb{C}$ *such that*

$$f(z) + 2f\left(\frac{1}{z}\right) + 3f\left(\frac{z}{z-1}\right) = z\,.$$

Solution. Let

$$h(z) = \frac{1}{z}\,, \quad g(z) = \frac{1}{1-z}\,.$$

Then

$$h^2(z) = z\,, \quad g^2(z) = \frac{z-1}{z}\,, \quad g^3(z) = z\,.$$

Also

$$hg(z) = 1-z\,, \quad gh(z) = \frac{z}{z-1}\,,$$

where we have shortened $h \circ g$ and $g \circ f$ into hg and gh, respectively.

The reader is encouraged to verify that $\{id, h, g, g^2, hg, gh\}$ is a group under composition and to construct the group table.

Substituting z by $h(z), g(z), g^2(z), hg(z)$, and $gh(z)$ in the defining equation, we get

$$\begin{aligned}
f\left(\frac{1}{z}\right) + 2f(z) + 3f\left(\frac{1}{1-z}\right) &= \frac{1}{z}\,,\\
f\left(\frac{1}{1-z}\right) + 2f(1-z) + 3f\left(\frac{1}{z}\right) &= \frac{1}{1-z}\,,\\
f\left(\frac{z-1}{z}\right) + 2f(\frac{z}{z-1}) + 3f(1-z) &= \frac{z-1}{z}\,,\\
f(1-z) + 2f\left(\frac{1}{1-z}\right) + 3f\left(\frac{z-1}{z}\right) &= 1-z\,,\\
f\left(\frac{z}{z-1}\right) + 2f\left(\frac{z-1}{z}\right) + 3f(z) &= \frac{z}{z-1}\,.
\end{aligned}$$

These, together with the given equation, form a system of six equations. Solving the system, we obtain

$$f(z) = \frac{1}{24}\left(2z + \frac{2}{z} + \frac{6}{z-1} + 3\right) = \frac{2z^3 + z^2 + 5z - 2}{24z(z-1)}\,. \qquad \square$$

Chapter 5

Cauchy's Equations

In this chapter we study four functional equations that are solved by the linear, power, exponential, and logarithmic functions. These equations were studied by Augustin Louis Cauchy, and since then they have formed the cornerstone of the theory. The derivation of the solutions of these functional equations serves as a vehicle to introduce the reader to the central ideas in the functional equations. It is imperative that these simple results and the methodology behind them are memorized (this can be done with minimal effort) as they appear often in functional problems.

5.1 First Cauchy Equation

Problem 5.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function of a continuous real variable that satisfies the functional relation*

$$f(x+y) \;=\; f(x) + f(y)\,, \qquad \forall x, y \in \mathbb{R}\,. \tag{5.1}$$

Find all functions f that satisfy the above conditions.

Comment. Equation (5.1) is known as the **first Cauchy functional equation** or the **linear Cauchy functional equation**.

Solution. Setting $x = y = 0$ in the defining equation, we see that

$$f(0) \;=\; 2f(0) \;\Rightarrow\; f(0) \;=\; 0\,.$$

Also, if $x = -y$,

$$f(0) \;=\; f(x) + f(-x) \;\Rightarrow\; f(-x) \;=\; -f(x)\,. \tag{5.2}$$

We now notice that for any natural number $n > 0$,

$$f(x_1 + x_2 + \cdots + x_n) \;=\; f(x_1) + f(x_2) + \cdots + f(x_n)\,. \tag{5.3}$$

This is verified using the method of induction. For $n = 2$, equation (5.3) is true by definition. Assuming that it is true for some n_0,

$$f(x_1 + x_2 + \cdots + x_{n_0}) = f(x_1) + f(x_2) + \cdots + f(x_{n_0}) ,$$

we see that it is true for $n_0 + 1$:

$$\begin{aligned} f(x_1 + x_2 + \cdots + x_{n_0} + x_{n_0+1}) &= f((x_1 + x_2 + \cdots + x_{n_0}) + x_{n_0+1}) \\ &= f(x_1 + x_2 + \cdots + x_{n_0}) + f(x_{n_0+1}) \\ &= f(x_1) + f(x_2) + \cdots + f(x_{n_0}) + f(x_{n_0+1}) . \end{aligned}$$

This proves the statement; that is, equation (5.3) is true for all natural numbers.

Setting $x_i = x,\ \forall i$ in (5.3),

$$f(nx) = n\, f(x) . \tag{5.4}$$

If moreover, $x = m\,z/n$, with $m \in \mathbb{N}$, $n \in \mathbb{N}^*$,

$$f(mz) = n\, f\left(\frac{m}{n} z\right) \overset{(5.4)}{\Rightarrow} \frac{m}{n}\, f(z) = f\left(\frac{m}{n} z\right) . \tag{5.5}$$

From this equation and (5.2), we conclude that

$$f\left(-\frac{m}{n} z\right) = -\frac{m}{n}\, f(z) . \tag{5.6}$$

In other words, so far we have proved that

$$f(qz) = q\, f(z) , \quad \forall z \in \mathbb{R} ,\ q \in \mathbb{Q} . \tag{5.7}$$

When $z = 1$,

$$f(q) = c\, q ,$$

where we set $c \equiv f(1)$.

Now, let $r \in \mathbb{R}$. There is a sequence of rational numbers (q_n) such that

$$\lim_{n \to +\infty} q_n = r .$$

For the terms of the sequence, the function f gives

$$f(q_n) = c\, q_n ,$$

and since it is continuous

$$f(r) = f\Big(\lim_{n \to +\infty} q_n\Big) = \lim_{n \to +\infty} f(q_n) = \lim_{n \to +\infty} c\, q_n = c\, r . \qquad \square$$

5.2 Second Cauchy Equation

Problem 5.2. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function of a real variable that satisfies the functional relation*

$$f(x+y) = f(x)\, f(y)\,, \quad \forall x, y \in \mathbb{R}\,. \tag{5.8}$$

Find all functions f that satisfy the above conditions and are not identically zero.

Comment. Equation (5.8) is known as the **second Cauchy functional equation** or the **exponential Cauchy functional equation**.

Solution 1. First, we will show that $f(x)$ is positive for any real number x. From the defining relation, we see that $f(x) = f(x/2)^2$, and therefore $f(x) \geq 0$. Let x_0 be a real number for which $f(x_0) = 0$. Then $f(x) = f(x - x_0 + x_0) = f(x - x_0)\, f(x_0) = 0$, i.e. the function would vanish identically. If there are solutions which do not vanish identically, then they cannot vanish at any point.

Since $f(x) > 0$, the function $g(x) = \ln f(x)$ is well-defined. Then the defining equation (5.8) can be written in terms of the function g as

$$g(x) + g(y) = g(x+y)\,, \quad \forall x, y \in \mathbb{R}\,,$$

which has the unique continuous solution $g(x) = c\, x$. From this the function f is found to be

$$f(x) = a^x\,,$$

where we set $a = e^c$. □

The above solution of course assumes that the result of the linear Cauchy functional equation (5.1) is known. However, one might want to establish a solution which is independent of the results for (5.1). This is done in the solution that follows.

Solution 2. One first proves that $f(x) \neq 0$ (as above). Setting $x = y = 0$ in the defining equation, we see that

$$f(0) = f(0)^2 \Rightarrow f(0)(f(0) - 1) = 0 \Rightarrow f(0) = 0,\ 1\,,$$

and therefore we must have $f(0) = 1$.

Also, if $x = -y$,

$$f(0) = f(x)\, f(-x) \Rightarrow f(-x) = f(x)^{-1}\,. \tag{5.9}$$

We now notice that for any natural number $n > 0$

$$f(x_1 + x_2 + \cdots + x_n) = f(x_1)\, f(x_2) \cdots f(x_n)\,. \tag{5.10}$$

This is verified using the method of induction. For $n = 2$, equation (5.10) is true by definition. Assuming, that it is true for some n_0,

$$f(x_1 + x_2 + \cdots + x_{n_0}) = f(x_1)\, f(x_2) \cdots f(x_{n_0})\,,$$

we see that it is true for $n_0 + 1$:

$$\begin{aligned} f(x_1 + x_2 + \cdots + x_{n_0} + x_{n_0+1}) &= f\Big((x_1 + x_2 + \cdots + x_{n_0}) + x_{n_0+1}\Big) \\ &= f(x_1 + x_2 + \cdots + x_{n_0})\, f(x_{n_0+1}) \\ &= f(x_1)\, f(x_2) \cdots f(x_{n_0})\, f(x_{n_0+1}) \,. \end{aligned}$$

Setting $x_i = x$, for all i in (5.10),

$$f(nx) = f(x)^n \,. \tag{5.11}$$

If moreover, $x = my/n$, $m \in \mathbb{N}$, $n \in \mathbb{N}^*$,

$$f(my) = f\left(\frac{my}{n}\right)^n \overset{(5.11)}{\Rightarrow} f(y)^m = f\left(\frac{m}{n}y\right)^n \Rightarrow f\left(\frac{m}{n}y\right) = f(y)^{m/n} \,. \tag{5.12}$$

From this equation and (5.9), we conclude that

$$f\left(-\frac{m}{n}y\right) = f(y)^{-m/n} \,. \tag{5.13}$$

In other words, so far we have proved that

$$f(qz) = f(z)^q \,, \quad \forall z \in \mathbb{R} \,, \; q \in \mathbb{Q} \,.$$

When $z = 1$,

$$f(q) = a^q \,,$$

where we set $a \equiv f(1)$.

Now, let $r \in \mathbb{R}$. There is a sequence of rationals $\{q_n\}$ such that

$$\lim_{n\to+\infty} q_n = r \,.$$

For the terms of the sequence $\{q_n\}$, the function f gives

$$f(q_n) = a^{q_n} \,,$$

and since it is continuous

$$f(r) = f\Big(\lim_{n\to+\infty} q_n\Big) = \lim_{n\to+\infty} f(q_n) = \lim_{n\to+\infty} a^{q_n} = a^r \,. \qquad \square$$

5.3 Third Cauchy Equation

Problem 5.3. *Let $f : \mathbb{R}^*_+ \to \mathbb{R}$ be a continuous function of a continuous real variable that satisfies the functional relation*

$$f(x) + f(y) = f(xy) \,, \quad \forall x, y \in \mathbb{R}^*_+ \,. \tag{5.14}$$

Find all functions f that satisfy the above conditions.

Comment. Equation (5.14) is known as the **third Cauchy functional equation** ot the **logarithmic Cauchy functional equation**.

Solution 1. Since the independent variable x takes values in $\mathbb{R}_+^*$, the change of variable $w = \ln x$ leads to a variable taking values in $\mathbb{R}$. Then the function $g(x) = f(e^x) \Leftrightarrow f(x) = g(\ln x)$ is a function $g : \mathbb{R} \to \mathbb{R}$, satisfying the relation

$$g(\ln x) + g(\ln y) \;=\; g\,(\ln(xy)) \;=\; g(\ln x + \ln y)\,,$$

or

$$g(w_1) + g(w_2) \;=\; g(w_1 + w_2)\,.$$

The solution to this functional relation is $g(w) = c\,w$ which can be inverted to give $f(x) = c\,\ln x = \ln x/\ln\gamma = \log_\gamma x$, where we set $c = 1/\ln\gamma$. □

The above solution of course assumes that we have worked out the exponential Cauchy functional equation (5.8). However, one might want to establish the result without any reference to that solution. This is done in the solution that follows.

Solution 2. Setting $x = y = 1$ in the defining equation (5.14), we see that

$$f(1) \;=\; 2f(1) \;\Rightarrow\; f(1) \;=\; 0\,.$$

Also, if $y = 1/x$,

$$f(1) \;=\; f(x) + f(x^{-1}) \;\Rightarrow\; f(x^{-1}) \;=\; -\,f(x)\,. \tag{5.15}$$

We now notice that for any natural number $n > 0$

$$f(x_1x_2\cdots x_n) \;=\; f(x_1) + f(x_2) + \cdots + f(x_n)\,. \tag{5.16}$$

This is verified using the method of induction. For $n = 2$, equation (5.16) is true by definition. Assuming, that it is true for some n_0,

$$f(x_1x_2\cdots x_{n_0}) \;=\; f(x_1) + f(x_2) + \cdots + f(x_{n_0})\,,$$

we see that it is true for $n_0 + 1$:

$$\begin{aligned} f(x_1x_2\cdots x_{n_0}x_{n_0+1}) &= f\Big((x_1x_2\cdots x_{n_0})x_{n_0+1}\Big) \\ &= f(x_1x_2\cdots x_{n_0})\,f(x_{n_0+1}) \\ &= f(x_1)\,f(x_2)\,\cdots\,f(x_{n_0})\,f(x_{n_0+1})\,. \end{aligned}$$

Setting $x_i = x$, for all i in (5.16),

$$f(x^n) \;=\; n\,f(x)\,. \tag{5.17}$$

If moreover, $x = y^{m/n}$, $m \in \mathbb{N}$, $n \in \mathbb{N}^*$,

$$f(y^m) \;=\; n\, f\left(y^{\frac{m}{n}}\right) \;\overset{(5.17)}{\Rightarrow}\; m\, f(y) \;=\; n\, f\left(y^{\frac{m}{n}}\right) \;\Rightarrow\; f\left(y^{\frac{m}{n}}\right) \;=\; \frac{m}{n}\, f(y)\,. \tag{5.18}$$

From this equation and (5.15), we conclude that

$$f\left(y^{-\frac{m}{n}}\right) \;=\; -\,\frac{m}{n}\, f(y)\,. \tag{5.19}$$

In other words, so far we have proved that

$$f(y^q) \;=\; q\, f(y)\,, \quad \forall y \in \mathbb{R}\,, \;\; q \in \mathbb{Q}\,.$$

Now, let $r \in \mathbb{R}$. There is a sequence of rationals $\{q_n\}$ such that

$$\lim_{n\to+\infty} q_n \;=\; r\,.$$

For the terms of the sequence $\{q_n\}$, the function f gives

$$f(y^{q_n}) \;=\; q_n\, f(y)\,,$$

and since it is continuous

$$f(y^r) \;=\; \lim_{n\to+\infty} f(y^{q_n}) \;=\; \lim_{n\to+\infty} q_n\, f(y) \;=\; r\, f(y)\,.$$

Finally, let's set $y = a = \text{const}$ and $x = a^r$. Then we find the function f in the final form:

$$f(x) \;=\; f(a)\, \log_a x \;=\; \log_b x\,,$$

where we defined a constant b such that $f(a) = 1/\log_a b$. □

5.4 Fourth Cauchy Equation

Problem 5.4. *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a continuous function of a real variable that satisfies the functional relation*

$$f(xy) \;=\; f(x)\, f(y)\,, \qquad \forall x, y \in \mathbb{R}_+\,. \tag{5.20}$$

Find all functions f that satisfy the above conditions and do not vanish identically.

Comment. Equation (5.20) is known as the **fourth Cauchy functional equation** or the **power Cauchy functional equation**.

Solution 1. First, let's show that $f(x) > 0$ for all $x \neq 0$. From the defining equation (5.20), we see that $f(x) = f(\sqrt{x})^2$. This implies that $f(x) \geq 0$, $\forall x \in \mathbb{R}_+$. Now, if there exists a point $x_0 \neq 0$ such that $f(x_0) = 0$, then

$$f(x) \;=\; f\left(\frac{x}{x_0}\, x_0\right) \;=\; f\left(\frac{x}{x_0}\right) f(x_0) \;=\; 0\,,$$

for all $x \in \mathbb{R}_+$. Therefore the function f would vanish identically. Since we are looking for non-vanishing functions f, there cannot be any x_0 with the above property. In other words, $f(x) > 0$, $\forall x \in \mathbb{R}_+$.

We now define $g(x) = \ln f(x)$. Then the defining relation (5.20), by taking the logarithm of the two sides, can be written in terms of the function g as:

$$g(x) + g(y) \;=\; g(x\,y)\,.$$

This relation has the solution $g(x) = 0$, or $g(x) = \log_\gamma x = \ln x/\ln\gamma$ and therefore

$$f(x) \;=\; 1\,, \quad \text{or} \quad f(x) \;=\; e^{g(x)} = x^c\,,$$

where $c = 1/\ln\gamma \neq 0$. □

The above solution of course assumes that we have worked out the logarithmic functional equation (5.14). However, one might want to establish the result without any reference to that solution. This is done in the solution that follows.

Solution 2. One first proves that $f(x) > 0$ (as above). Setting $x = y = 1$ in the defining equation, we see that

$$f(1) \;=\; f(1)^2 \;\Rightarrow\; f(1)\,(f(1) - 1) \;=\; 0 \;\Rightarrow\; f(1) \;=\; 0,\, 1\,,$$

and therefore we must have $f(1) = 1$.

Also, if $y = 1/x$ for $x \neq 0$,

$$f(1) \;=\; f(x)\,f(x^{-1}) \;\Rightarrow\; f(x^{-1}) \;=\; f(x)^{-1}\,. \tag{5.21}$$

We now notice that for any natural number $n > 0$

$$f(x_1 x_2 \cdots x_n) \;=\; f(x_1)\,f(x_2)\,\cdots\,f(x_n)\,. \tag{5.22}$$

This is verified using the method of induction. For $n = 2$, equation (5.22) is true by definition. Assuming, that it is true for some n_0,

$$f(x_1 x_2 \cdots x_{n_0}) \;=\; f(x_1)\,f(x_2)\,\cdots\,f(x_{n_0})\,,$$

we see that it is true for $n_0 + 1$:

$$\begin{aligned} f(x_1 x_2 \cdots x_{n_0} x_{n_0+1}) \;&=\; f\Big((x_1 x_2 \cdots x_{n_0}) x_{n_0+1}\Big) \\ &=\; f(x_1 x_2 \cdots x_{n_0})\,f(x_{n_0+1}) \\ &=\; f(x_1)\,f(x_2)\,\cdots\,f(x_{n_0})\,f(x_{n_0+1})\,. \end{aligned}$$

Setting $x_i = x,\ \forall i$ in (5.22),

$$f(x^n) \;=\; f(x)^n\;. \tag{5.23}$$

If moreover, $x = y^{m/n}$, $m \in \mathbb{N}$, $n \in \mathbb{N}^*$,

$$f(y^m) \;=\; f\left(y^{\frac{m}{n}}\right)^n \overset{(5.23)}{\Rightarrow} f(y)^m \;=\; f\left(y^{\frac{m}{n}}\right)^n \;\Rightarrow\; f\left(y^{\frac{m}{n}}\right) \;=\; f(y)^{m/n}\;. \tag{5.24}$$

From this equation and (5.21), we conclude that

$$f\left(y^{-\frac{m}{n}}\right) \;=\; f(y)^{-m/n}\;. \tag{5.25}$$

In other words, so far we have proved that

$$f(y^q) \;=\; f(y)^q\;, \quad \forall y \in \mathbb{R}_+^*\;,\;\; q \in \mathbb{Q}\;.$$

Now, let $r \in \mathbb{R}$. There is a sequence of rationals $\{q_n\}$ such that

$$\lim_{n\to\infty} q_n \;=\; r\;.$$

For the terms of the sequence $\{q_n\}$, the function f gives

$$f(y^{q_n}) \;=\; f(y)^{q_n}\;,$$

and since it is continuous

$$f(y^r) \;=\; \lim_{n\to\infty} f(y^{q_n}) \;=\; \lim_{n\to\infty} f(y)^{q_n} \;=\; f(y)^r\;.$$

Finally, let's set $y = a =$ const and $x = a^r$. Then we find the function f in the form:

$$f(x) \;=\; f(a)^{\log_a x}\;.$$

If $f(a) = 1$, then $f(x) = 1$. If $f(a) \neq 1$, we define a constant c such that $f(a) = a^c$. Then

$$f(x) \;=\; x^c\;, \quad c \neq 0\;. \qquad \square$$

Question. *Define the sign function sgn(x) by*

$$sgn(x) \;=\; \begin{cases} -1\;, & \text{if } x < 0\;,\\ 0\;, & \text{if } x = 0\;,\\ +1\;, & \text{if } x > 0\;,\end{cases}$$

Show that it satisfies the equation

$$sgn(x\,y) \;=\; sgn(x)\,sgn(y)\;.$$

In other words it is a solution of the functional equation (5.20). However, it is not a power function. Explain.

Answer. The sign function is not continuous. It has a jump discontinuity at $x = 0$. When we presented the definition of functional equations in Section 2.1, we mentioned that the set of solutions strongly depends on the conditions imposed on the functions. When continuity is assumed, the Cauchy functional equations have a restricted set of solutions, exactly as we have seen above. If this condition is relaxed, then the set of solutions greatly enlarges. Non-continuous solutions of the Cauchy equations will be studied in Chapter 11.

5.5 Solved Problems

Often functional equations can be solved by a reduction to one of Cauchy's functional equations. In this section, we present several examples.

Problem 5.5 ([45]). *Find the continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional relation*

$$f(x+y) \;=\; A^y f(x) + A^x f(y)\,, \quad \forall x, y \in \mathbb{R}\,,$$

where A is a positive constant.

Solution. We define the continuous function $g(x) = A^{-x} f(x)$. Then the functional equation takes the form

$$g(x+y) \;=\; g(x) + g(y)\,,$$

which has the solution $g(x) = c\,x$ where c a constant. Therefore $f(x) = c\,x\,A^x$. □

Problem 5.6. *Find the continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f\left(\frac{x+y}{2}\right) \;=\; \frac{f(x)+f(y)}{2}\,, \quad \forall x, y \in \mathbb{R}\,. \tag{5.26}$$

Comment. Equation (5.26) is known as the **Jensen functional equation**. It can be equivalently written as

$$f(x+y) + f(x-y) \;=\; 2\,f(x)\,.$$ □

Solution. For $y = 0$, equation (5.26) gives

$$f\left(\frac{x}{2}\right) \;=\; \frac{f(x)+f(0)}{2} \;=\; \frac{f(x)+b}{2}\,,$$

where $b = f(0)$. In this relation we substitute $x = y + z$:

$$f\left(\frac{y+z}{2}\right) \;=\; \frac{f(y+z)+b}{2}\,.$$

However, the left hand side can be rewritten from the defining relation (5.26) and thus

$$\frac{f(y)+f(z)}{2} \;=\; \frac{f(y+z)+b}{2}\,,$$

or

$$f(y+z) \;=\; f(y) + f(z) - b\,.$$

We define the continuous function $g(x) = f(x) - b$. Then the functional equation takes the form

$$g(x+y) \;=\; g(x) + g(y)\,,$$

which has the solution $g(x) = c\,x$ where c is a constant. Therefore $f(x) = c\,x + b$. □

Problem 5.7. *Find the continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f\left(\sqrt{\frac{x^2+y^2}{2}}\right) = \sqrt{\frac{f(x)^2+f(y)^2}{2}}\,, \quad \forall x, y \in \mathbb{R}\,.$$

Solution. From the defining equation we see that $f(x) \geq 0$, $\forall x \in \mathbb{R}$ (since the square root of a number is a non-negative number). Even more,

$$\sqrt{\frac{f(x)^2+f(y)^2}{2}} = f\left(\sqrt{\frac{x^2+y^2}{2}}\right) = \sqrt{\frac{f(-x)^2+f(y)^2}{2}}$$
$$\Rightarrow f(x)^2 = f(-x)^2 \Rightarrow f(x) = f(-x)\,,$$

i.e. the function f is even.

Let's define the function $F(x) = f(\sqrt{x})^2$ for $x \geq 0$. Then the defining equation becomes

$$F\left(\frac{u+v}{2}\right) = \frac{F(u)+F(v)}{2}\,,$$

where $u = x^2$, $v = y^2$. Therefore,

$$F(u) = c\,u + b\,,$$

or

$$f(x) = \sqrt{c\,x^2+b}\,.$$ □

Problem 5.8. *Find the continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(x+y) = a^{xy}\,f(x)\,f(y)\,, \quad \forall x, y \in \mathbb{R}\,,$$

where a is a positive constant.

Solution. We define the continuous function $g(x) = a^{-x^2/2}\,f(x)$. Then the functional equation takes the form

$$g(x+y) = g(x)\,g(y)\,,$$

which has the solutions $g(x) = 0$ or $g(x) = b^x$ where b is a positive constant. Therefore $f(x) = 0$ or $f(x) = b^x\,a^{x^2/2}$. □

Problem 5.9. *Find the continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(x+y) \;=\; f(x) + f(y) + f(x)\,f(y)\,, \quad \forall x, y \in \mathbb{R}\,.$$

Solution. We define the continuous function $g(x) = f(x) + 1$. Then the functional equation takes the form

$$g(x+y) \;=\; g(x)\,g(y)\,,$$

which has the solutions $g(x) = 0$ or $g(x) = b^x$ where b a positive constant. Therefore $f(x) = -1$ or $f(x) = b^x - 1$. □

Problem 5.10 (Putnam 1947)**.** *Find the continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(\sqrt{x^2 + y^2}) \;=\; f(x)\,f(y)\,, \quad \forall x, y \in \mathbb{R}\,.$$

Solution. If f is vanishing identically, it satisfies the equation. Therefore, we shall look for solutions that do not vanish identically. This implies that there exists a real number x_0 such that $f(x_0) \neq 0$ and then $f(\sqrt{x_0^2 + y^2}) = f(x_0)\,f(y) = f(x_0)\,f(-y)$, or $f(y) = f(-y)$, for all $y \in \mathbb{R}$. The function f is necessarily even. It is enough to search for the form of $f(x)$ when $x \geq 0$.

We define the continuous function $g(x) = f(\sqrt{x})$, $x \geq 0$. Then the functional equation takes the Cauchy form

$$g(u+v) \;=\; g(u)\,g(v)\,,$$

where $u = x^2$, $v = y^2$. This equation has the non-vanishing solution $g(u) = a^u$ where a is a positive constant. Therefore $f(x) = a^{x^2}$. □

Problem 5.11. *Find the continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(\sqrt[n]{x^n + y^n}) \;=\; f(x) + f(y)\,, \quad \forall x, y \in \mathbb{R}\,,$$

where n is a given positive natural number.

Solution. We notice that if $n = 2k$, then $f(\sqrt[n]{x^n + y^n}) = f(x) + f(y) = f(-x) + f(y)$, or $f(x) = f(-x)$, for all $x \in \mathbb{R}$. We define the continuous function $g(x) = f(\sqrt[n]{x})$, $x \geq 0$, if $n = 2k$ and $g(x) = f(\sqrt[n]{x})$, if $n = 2k+1$. Then the functional equation takes the Cauchy form

$$g(u+v) \;=\; g(u) + g(v)\,,$$

where $u = x^n$, $v = y^n$. This equation has the solution $g(u) = a\,u$ where a is a constant. Therefore $f(x) = a\,x^n$. □

Problem 5.12 (Romania 1997). *Find all continuous solutions $f : \mathbb{R} \to [0, +\infty)$ such that*

$$f(x^2 + y^2) = f(x^2 - y^2) + f(2xy) .$$

Solution. For $x = y = 0$, the defining equation gives $f(0) = 0$. Then for $x = 0$ only, $f(y^2) = f(-y^2)$. That is, the function f is even, and we only need to find its value for positive reals.

For $x > y$, we define

$$\alpha = x^2 - y^2 , \quad \beta = 2xy .$$

Then $\alpha^2 + \beta^2 = (x^2 + y^2)^2$ and the given functional equation takes the form

$$f(\sqrt{\alpha^2 + \beta^2}) = f(\alpha) + f(\beta) .$$

Using the previous results, $f(x) = a\,x^2$ where a is a real constant. □

Problem 5.13. *Find the continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(x + y) = x^2 y + xy^2 - 2xy + f(x) + f(y) , \quad \forall x, y \in \mathbb{R} .$$

Solution. We define the continuous function $g(x) = f(x) - x^3/3 + x^2$. Then the functional equation takes the Cauchy form

$$g(x + y) = g(x) + g(y) ,$$

This equation has the solution $g(x) = a\,x$ where a is a constant. Therefore $f(x) = a\,x + x^3/3 - x^2$. □

Problem 5.14. *Show that the functional equation*

$$f\left(\frac{x + y}{2}\right)^2 = f(x)\, f(y) \tag{5.27}$$

is equivalent to the functional equation

$$f(x)^2 = f(x + y)\, f(x - y) . \tag{5.28}$$

Then find the continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of (5.27).

Comment. Equation (5.28) is known as the **Lobacevskii functional equation**. □

Solution. Setting $x = a + b$ and $y = a - b$ in (5.27), we find (5.28).

Setting $y = 0$ in (5.27), we find

$$f\left(\frac{x}{2}\right)^2 = f(x)\, f(0)\,.$$

If $f(0) = 0$, then $f(x/2)^2 = 0$, $\forall x \in \mathbb{R}$. This implies that $f(x) = 0$, $\forall x \in \mathbb{R}$. If $f(0) \neq 0$, we replace x with $x + y$, then:

$$f\left(\frac{x+y}{2}\right)^2 = f(x+y)\, f(0)\,,$$

or, after using (5.27):

$$f(x)\, f(y) = f(x+y)\, f(0)\,.$$

We define the continuous function $g(x) = f(x)/f(0)$. Then the functional equation takes the Cauchy form

$$g(x+y) = g(x)\, g(y)\,.$$

This equation has the solutions $g(x) = 0$ or $g(x) = a^x$ where $a \neq 0$; therefore $f(x) = f(0)\, a^x$. □

Problem 5.15 ([45]). *Find all continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(x+y) = f(x) + f(y) + a\,(1 - b^x)\,(1 - b^y)\,, \tag{5.29}$$

where a, b are real constants and $b > 0$.

Solution. Set

$$g(x) = f(x) - a\,(b^x - 1)\,. \tag{5.30}$$

Then the function $g(x)$ satisfies

$$g(x+y) = g(x) + g(y)\,,$$

with solution $g(x) = c\,x$ with c a real constant. Therefore $f(x) = a\,(b^x - 1) + c\,x$. □

Problem 5.16. *Let $a, b \in \mathbb{R}$. Find the continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(x+y+a) + b = f(x) + f(y)\,, \quad \forall x, y \in \mathbb{R}\,. \tag{5.31}$$

Solution. Define the continuous function $g(x) = f(x - a) - b$. Then (5.31) takes the form

$$g(x + y + 2a) = g(x + a) + g(y + a) ,$$

or

$$g(u + v) = g(u) + g(v) ,$$

where $u = x + a$ and $v = y + a$. By (5.1), $g(u) = cu$, where c is a constant. Therefore $f(x) = c(x + a) + b$. □

Problem 5.17 (Romania TST 2006). *Let $r, s \in \mathbb{Q}$. Find all functions $f : \mathbb{Q} \to \mathbb{Q}$ such that for all $x, y \in \mathbb{Q}$ we have*

$$f(x + f(y)) = f(x + r) + y + s . \tag{5.32}$$

Solution. It is easy to verify that $f(x) = \pm(x + s) + r$ are solutions. We show that they are the only solutions.

Adding $z - r$ to both sides of (5.32) and finding the value of f at the points thus obtained, we get

$$f(z - r + f(x + f(y))) = f(z + y + s - r + f(x + r)) . \tag{5.33}$$

Applying (5.32) to each side of (5.33), we obtain

$$f(z) + x + f(y) + s = f(z + y + s) + x + r + s ,$$

or

$$f(y) + f(z) = f(y + z + s) + r ,$$

which is (5.31). Therefore $f(x) = c(x+s)+r$ where c is a constant. Substituting $f(x) = c(x+s)+r$ into (5.32), we find

$$c(x + c(y + s) + r + s) + r = c(x + r + s) + r + y + s ,$$

or

$$(c^2 - 1)(y + s) = 0 .$$

Hence $c = \pm 1$ and $f(x) = \pm(x + s) + r$. □

Comment. Similarly, if $a, b \in \mathbb{R}$, then $f(x) = \pm(x + a) + b$ are the only continuous solutions to the functional equation

$$f(x + f(y)) = f(x + a) + y + b , \quad \forall x, y \in \mathbb{R} .$$

The explanation is left as an easy exercise to the reader. □

Chapter 6

Cauchy's $\mathbb{NQR}$-Method

We notice that the functional equations (5.1) and (5.8) are of the form

$$f(x+y) = F\left(f(x), f(y)\right) \ , \tag{6.1}$$

where the function $F(a,b)$ is $F(a,b) = a+b$ or $F(a,b) = a\,b$. One can explore functional equations with different functions $F(a,b)$. For example,

$$\begin{aligned} f(x+y) &= f(x) + f(y) + f(x)f(y) \ , \\ f(x+y) &= \frac{f(x)+f(y)}{1-f(x)\,f(y)} \ , \end{aligned}$$

are functional equations of the form (6.1) with

$$\begin{aligned} F(a,b) &= a+b+ab \ , \\ F(a,b) &= \frac{a+b}{1-ab} \ , \end{aligned}$$

respectively.

One can generalize equation (6.1) further to

$$f(x+y) = F\left(f(x), f(y), f(x-y); x, y\right) \ .$$

Such examples are

$$\begin{aligned} f(x+y) &= k^{xy}\, f(x)f(y) \ , \\ f(x+y)f(x-y) &= f(x)^2 \ , \\ f(x+y)f(x-y) &= f(x)^2\, f(y)^2 \ , \\ f(x+y) + f(x-y) &= 2\,f(x) \ , \\ f(x+y) - f(x-y) &= 2\,f(y) \ , \\ f(x+y) + f(x-y) &= 2\,f(x) + 2\,f(y) \ , \\ f(x+y) + f(x-y) &= 2\,f(x)\cos y \ , \end{aligned}$$

with

$$\begin{aligned}
F(a,b,c;x,y) &= k^{xy}\,a\,b\,, \\
F(a,b,c;x,y) &= \frac{a^2}{c}\,, \\
F(a,b,c;x,y) &= \frac{a^2b^2}{c}\,, \\
F(a,b,c;x,y) &= 2\,a - c\,, \\
F(a,b,c;x,y) &= 2\,b + c\,, \\
F(a,b,c;x,y) &= 2\,a + 2\,b - c\,, \\
F(a,b,c;x,y) &= 2a\,\cos y - c\,,
\end{aligned}$$

respectively.

To include the functional relations (5.14) and (5.20), we must generalize the equation further to allow arbitrary functions of x and y in the left hand side and not only the sum $x+y$:

$$f\,(G(x,y)) = F\,(f(x), f(y), f(x-y); x, y)\ .$$

Such examples are

$$\begin{aligned}
f(x\,y) &= f(x) + f(y) + f(x)f(y)\,, \\
f(\sqrt[n]{x^n + y^n}) &= f(x) + f(y)\,, \\
f(\sqrt[n]{x^n + y^n}) &= f(x)\,f(y)\,,
\end{aligned}$$

with

$$\begin{aligned}
G(x,y) &= x\,y\,, \quad F(a,b,c;x,y) = a+b+ab\,, \\
G(x,y) &= \sqrt[n]{x^n + y^n}\,, \quad F(a,b,c;x,y) = a+b\,, \\
G(x,y) &= \sqrt[n]{x^n + y^n}\,, \quad F(a,b,c;x,y) = a\,b\,,
\end{aligned}$$

respectively.

For such functional relations, often, the method in which we first reconstruct the function in $\mathbb{N}$, then in $\mathbb{Q}$, and finally in $\mathbb{R}$ is successful. This method was demonstrated in Chapter 5 for the solution of (5.1), (5.8) (5.14), and (5.20). We shall name it the $\mathbb{NQR}$ method[1].

6.1 The $\mathbb{NQR}$-method

We present now the sequence of steps to obtain the continuous solutions of the equation

$$f(x+y) = F\,(f(x), f(y), f(x-y); x, y)\ . \tag{6.2}$$

[1]This name was made by the author; it is not encountered in the literature of the topic.

In this more general case, we follow the same steps; however, the operation of summation in the left hand side has to be changed to an operation which will be dictated by the function $G(x, y)$.

First of all, we set $x = y = 0$ to find an equation for $f(0)$:

$$f(0) = F\left(f(0), f(0), f(0); 0, 0\right) .$$

The value of f at $x = 0$ is then known by solving this equation.

In (6.2), we substitute successively x for $2x, 3x, \ldots, (n-1)x$ and $y = x$:

$$\begin{aligned}
f(2x) &= F\left(f(x), f(x), f(0); x, x\right) \equiv F_2\left(f(x), x\right) , \\
f(3x) &= F\left(f(2x), f(x), f(x); 2x, x\right) \equiv F_3\left(f(x), x\right) , \\
f(4x) &= F\left(f(3x), f(x), f(2x); 3x, x\right) \equiv F_4\left(f(x), x\right) , \\
\ldots &\quad \ldots \\
f(nx) &= F\left(f\left((n-1)x\right), f(x), f\left((n-2)x\right); (n-1)x, x\right) \equiv F_n\left(f(x), x\right) .
\end{aligned}$$

In the last equation, we substitute $x = my/n$

$$f(my) = F_n\left(f\left(\frac{m}{n}y\right), \frac{m}{n}y\right) .$$

The left hand side however may be rewritten using the previous result:

$$F_m\left(f(y), y\right) = F_n\left(f\left(\frac{m}{n}y\right), \frac{m}{n}y\right) .$$

We assume that this expression can be solved for $f(\frac{m}{n}y)$ to give

$$f\left(\frac{m}{n}y\right) = \mathcal{F}\left(f(y), y; \frac{m}{n}\right) ,$$

for some function $\mathcal{F}(a, y; x)$. For $y = 1$,

$$f(q) = \mathcal{F}\left(f(1), 1; q\right) , \quad \forall q \in \mathbb{Q}_+ .$$

Let x be any non-negative real number. Then, there is a sequence of non-negative rational numbers $\{q_n\}$ which converges to x. Assuming continuity of f and $\mathcal{F}$, the last expression gives

$$f(x) = \mathcal{F}\left(f(1), 1; x\right) , \quad \forall x \in \mathbb{R}_+ .$$

To find the function $f(x)$ for $x < 0$, we relate it to $f(-x)$ using the defining equation (6.2). In particular, setting $y = -x$ in (6.2) we find:

$$\begin{aligned}
f(0) &= F\left(f(x), f(-x), f(2x); x, -x\right) \\
&= F\left(f(x), \mathcal{F}\left(f(1), 1; -x\right), F_2\left(f(x), x\right); x, -x\right) .
\end{aligned}$$

This should be solved for $f(x)$ to find the final result:

$$f(x) = \mathcal{G}(x) , \quad x < 0 .$$

6.2 Solved Problems

Problem 6.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function that satisfies the functional relation*

$$f(x+y) = f(x) + f(y) + f(x)\,f(y)\,, \qquad \forall x, y \in \mathbb{R}\,.$$

Find all functions f that satisfy the above conditions.

Solution. Setting now $x = y = z/2$ in the defining equation, we see that

$$f(z) = 2f\left(\frac{z}{2}\right) + f\left(\frac{z}{2}\right)^2 = \left(1 + f\left(\frac{z}{2}\right)\right)^2 - 1\,,$$

i.e. $f(x) \geq -1$, $\forall x \in \mathbb{R}$. In particular, $f(1) \geq -1$. If $f(1) = -1$, then $f(x) = f(x-1+1) = f(x-1) + f(1) + f(x-1)\,f(1) = -1$, $\forall x \in \mathbb{R}$. We shall search for non-constant solutions, and thus $f(1) > -1$.

Setting $y = x, 2x, 3x, \ldots$, in the defining relation, we find

$$\begin{aligned} f(2x) &= 2f(x) + f(x)^2 = (f(x)+1)^2 - 1\,, \\ f(3x) &= f(2x) + f(x) + f(2x)\,f(x) = f(x)^3 + 3f(x)^3 + 3f(x) = (f(x)+1)^3 - 1\,, \\ \cdots & \quad \cdots\,, \end{aligned}$$

which indicates that

$$f(nx) = (f(x)+1)^n - 1\,, \quad n \in \mathbb{N}\,. \tag{6.3}$$

We can prove this formula for any $n \in \mathbb{N}$ using induction. For $m = 1, 2$ it is true. Let's assume that it is true for $n = m$:

$$f(mx) = (f(x)+1)^m - 1\,. \tag{6.4}$$

Then we shall show that it is also true for $n = m+1$. In the defining relation, we set $y = mx$:

$$f((m+1)x) = f(x) + f(mx) + f(x)\,f(mx)\,.$$

From (6.4) we can compute the right hand side:

$$f((m+1)x) = f(x) + (f(x)+1)^m - 1 + f(x)\,(f(x)+1)^m - f(x) = (f(x)+1)^{m+1} - 1\,.$$

This completes the induction.

Now we substitute $x = my/n$, $m \in \mathbb{N}$, $n \in \mathbb{N}^*$, in (6.3):

$$f(my) = \left(f\left(\frac{m}{n}y\right) + 1\right)^n - 1\,,$$

or

$$f\left(\frac{m}{n}y\right) = (f(y)+1)^{\frac{m}{n}} - 1\,.$$

For $y = 1$:

$$f(q) = c^q - 1\,, \quad \forall q \in \mathbb{Q}_+\,,$$

where $c = 1 + f(1)$.

Now, let $r \in \mathbb{R}_+$. There is a sequence of rationals $\{q_n\}$ such that

$$\lim_{n\to\infty} q_n = r\,.$$

For the terms of the sequence $\{q_n\}$, the function f gives

$$f(q_n) = c^{q_n} - 1\,,$$

and since f is continuous

$$f(r) = f\Big(\lim_{n\to\infty} q_n\Big) = \lim_{n\to\infty} f(q_n) = \lim_{n\to\infty} c^{q_n} - 1 = c^r - 1\,,$$

$\forall r \in \mathbb{R}_+$.

For $r = 0$, $f(0) = 0$. If $y = -x$ with $x > 0$ in the defining equation,

$$f(0) = f(x) + f(-x) + f(x)\,f(-x) \Rightarrow f(-x) = \frac{-f(x)}{1 + f(x)} = c^{-x} - 1\,.$$

Therefore,

$$f(x) = c^x - 1\,, \quad \forall x \in \mathbb{R}\,.$$

□

Problem 6.2. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function that satisfies the functional relation*

$$f(x + y) = a^{xy}\, f(x)\, f(y)\,, \qquad \forall x, y \in \mathbb{R}\,,$$

where a is a positive real constant. Find all functions f that satisfy the above conditions and are not identically zero.

Solution. Let x_0 be a real number for which $f(x_0) = 0$. Then $f(x) = f(x - x_0 + x_0) = a^{(x-x_0)x_0}\, f(x - x_0)\, f(x_0) = 0$, i.e. the function would vanish identically. If there are solutions which do not vanish identically, then f cannot vanish at any point. Setting now $x = y = 0$ in the defining equation, we see that

$$f(0) = f(0)^2 \Rightarrow f(0)\,(f(0) - 1) = 0 \Rightarrow f(0) = 0,\ 1\,,$$

and therefore we must have $f(0) = 1$.

Setting $y = x, 2x, \ldots, (n-1)x$ in the defining expression, we find

$$\begin{aligned} f(2x) &= f(x+x) = a^{x^2} f(x)^2 , \\ f(3x) &= f(x+2x) = a^{(1+2)x^2} f(x)^3 = a^{3x^2} f(x)^3 , \\ f(4x) &= f(x+3x) = a^{(1+2+3)x^2} f(x)^4 = a^{6x^2} f(x)^4 , \\ \ldots & \quad \ldots \\ f(nx) &= f\left(x+(n-1)x\right) = a^{(1+2+3+\cdots+(n-1))x^2} f(x)^n = a^{\frac{n(n-1)}{2}x^2} f(x)^n . \end{aligned}$$

If moreover, $x = my/n$, $m \in \mathbb{N}$, $n \in \mathbb{N}^*$,

$$f(my) = a^{\frac{(n-1)}{2n}m^2y^2} f\left(\frac{m}{n}y\right)^n ,$$

or

$$a^{\frac{m(m-1)}{2}y^2} f(y)^m = a^{\frac{(n-1)}{2n}m^2y^2} f\left(\frac{m}{n}y\right)^n .$$

This expression can be solved easily for $f(\frac{m}{n}y)$:

$$f\left(\frac{m}{n}y\right) = a^{\frac{1}{2}\frac{m}{n}(\frac{m}{n}-1)y^2} f(y)^{m/n} .$$

For $y = 1$:

$$f(q) = a^{\frac{q(q-1)}{2}} c^q , \quad \forall q \in \mathbb{Q}_+ ,$$

where $c = f(1)$.

Now, let $r \in \mathbb{R}_+$. There is a sequence of rationals $\{q_n\}$ such that

$$\lim_{n\to\infty} q_n = r .$$

For the terms of the sequence $\{q_n\}$, the function f gives

$$f(q_n) = a^{\frac{q_n(q_n-1)}{2}} c^{q_n} ,$$

and since f is continuous

$$f(r) = f\Big(\lim_{n\to+\infty} q_n\Big) = \lim_{n\to+\infty} f(q_n) = \lim_{n\to+\infty} a^{\frac{q_n(q_n-1)}{2}} c^{q_n} = a^{\frac{r(r-1)}{2}} c^r ,$$

for all $r \in \mathbb{R}_+$.

Now, if $y = -x$, $x > 0$ in the defining equation,

$$f(0) = a^{-x^2} f(x) f(-x) \Rightarrow f(-x) = a^{x^2} f(x)^{-1} = a^{\frac{-x(-x-1)}{2}} c^{-x} .$$

Therefore,

$$f(x) = a^{\frac{x(x-1)}{2}} c^x , \quad \forall x \in \mathbb{R} . \qquad \square$$

Problem 6.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function that satisfies the functional relation*

$$f(x+y)\, f(x-y) \;=\; f(x)^2\, f(y)^2\,, \qquad \forall x, y \in \mathbb{R}\,.$$

Find all functions f that satisfy the above conditions and are not identically zero.

Solution. Setting now $x = y = 0$ in the defining equation, we see that

$$f(0)^4 \;=\; f(0)^2 \;\Rightarrow\; f(0)^2 \left(f(0)^2 - 1\right) \;=\; 0 \;\Rightarrow\; f(0) \;=\; 0,\; \pm 1\,.$$

Now let $x = y$:

$$f(2x)\, f(0) \;=\; f(x)^4\,, \qquad \forall x\,. \tag{6.5}$$

If $f(0) = 0$, then $f(x) = 0$, $\forall x$. Since we are looking for solutions that do not vanish identically, we shall assume $f(0) \neq 0$. Consider first that $f(0) = 1$.

Replacing x by $x, 2x, 2^2x, \ldots$, in (6.5), we find

$$\begin{aligned} f(2x) &= f(x)^4 \;=\; f(x)^{2^2}\,, \\ f(4x) &= f(2x)^4 \;=\; f(x)^{4^2}\,, \\ f(8x) &= f(4x)^4 \;=\; f(x)^{8^2}\,, \\ \cdots & \quad \cdots \end{aligned}$$

This indicates that

$$f(nx) \;=\; f(x)^{n^2}\,, \quad n \in \mathbb{N}\,. \tag{6.6}$$

We can prove this formula for any $n \in \mathbb{N}$ using induction. For $m = 1, 2$ it is true. Let's assume that it is true for $\forall n \leq m$:

$$f(nx) \;=\; f(x)^{n^2}\,, \quad n = 1, 2, \ldots, m\,. \tag{6.7}$$

Then we shall show that it is also true for $n = m + 1$. In the defining relation, we set $x = my$:

$$f\left((m+1)y\right)\, f\left((m-1)y\right) \;=\; f(my)^2\, f(y)^2\,.$$

We can solve this equation easily for $f((m+1)y)$ and use (6.7) to find:

$$f\left((m+1)y\right) \;=\; f(y)^{-(m-1)^2}\, f(y)^{2m^2}\, f(y)^2 \;=\; f(y)^{(m+1)^2}\,.$$

This completes the proof.

Now we substitute $x = my/n$, $m \in \mathbb{N}$, $n \in \mathbb{N}^*$, in (6.6):

$$f(my) \;=\; f\left(\frac{m}{n}y\right)^{n^2}\,,$$

or

$$f(y)^{m^2} = f\left(\frac{m}{n}y\right)^{n^2} .$$

This expression can be solved easily for $f(\frac{m}{n}y)$:

$$f\left(\frac{m}{n}y\right) = f(y)^{\frac{m^2}{n^2}} .$$

For $y = 1$:

$$f(q) = c^{q^2} , \quad \forall q \in \mathbb{Q}_+ ,$$

where $c = f(1)$.

Now, let $r \in \mathbb{R}_+$. There is a sequence of rationals $\{q_n\}$ such that

$$\lim_{n\to\infty} q_n = r .$$

For the terms of the sequence $\{q_n\}$, the function f gives

$$f(q_n) = c^{q_n^2} ,$$

and since f is continuous

$$f(r) = f\left(\lim_{n\to\infty} q_n\right) = \lim_{n\to\infty} f(q_n) = \lim_{n\to\infty} c^{q_n^2} = c^{r^2} ,$$

for all $r \in \mathbb{R}_+$.

Now, if $x = 0$ in the defining equation and rename y as x,

$$f(x)^2 f(0)^2 = f(x)\, f(-x) \Rightarrow f(-x) = f(x) .$$

Therefore,

$$f(x) = c^{x^2} , \quad \forall x \in \mathbb{R} .$$

Finally, if $f(0) = -1$ then $g(x) = -f(x)$ satisfies $g(x+y)g(x-y) = g(x)^2 g(y)^2$, $\forall x, y \in \mathbb{R}$ and $g(0) = 1$. Applying the argument above to g we find $g(x) = c^{x^2}$. Hence $f(x) = -c^{x^2}$, $\forall x \in \mathbb{R}$. □

The next problem is of the form $f(x+y) = F(f(x), f(y), f(xy))$ where $F(a, b, c) = a+b-ab+c$. It does not quite fit into the ℕℚℝ-method above. But a combination of an algebraic system and ℕℚℝ-idea still proves successful.

Problem 6.4 (India TST 2003)**.** *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that for all $x, y \in \mathbb{R}$,*

$$f(x+y) + f(x)\, f(y) = f(x) + f(y) + f(xy) . \tag{6.8}$$

Solution. Clearly, $f \equiv 0$, $f \equiv 2$, and $f(x) = x$ are solutions. We show that they are the only solutions.

Setting $x = 0 = y$ in (6.8) we get $f(0)^2 = 2f(0)$. If $f(0) = 2$, then by letting $y = 0$ in (6.8) we see that $f(x) = 2$, $\forall x \in \mathbb{R}$, i.e. $f \equiv 2$.

Suppose therefore that $f(0) = 0$. Let $a = f(1)$. Setting $x = 1$ and $y = -1$ we get

$$a\,f(-1) \;=\; a + 2\,f(-1) \;\Rightarrow\; f(-1) \;=\; \frac{a}{a-2}\,.$$

Now substituting (x, y) in (6.8) by $(x - 1, 1)$, $(-x + 1, -1)$, and $(-x, 1)$, we get

$$f(x) + (a-2)f(x-1) \;=\; a\,, \tag{6.9}$$

$$f(-x) + \left(\frac{2}{a-2}\right)f(-x+1) - f(x-1) \;=\; \frac{a}{a-2}\,, \tag{6.10}$$

$$f(-x+1) + (a-2)f(-x) \;=\; a\,. \tag{6.11}$$

Eliminating $f(x - 1)$ and $f(-x + 1)$ in (6.9)–(6.11), we get

$$f(x) - (a-2)f(-x) \;=\; 0\,. \tag{6.12}$$

Replacing x by $-x$ in (6.12) we get

$$f(-x) - (a-2)f(x) \;=\; 0\,. \tag{6.13}$$

If $a \notin \{1, 3\}$, then by eliminating $f(-x)$ in (6.12) and (6.13) we see that $f(x) = 0$, $\forall x \in \mathbb{R}$, i.e. $f \equiv 0$.

Consider next that $a = 3$. Then (6.9) gives $f(x) = 3 - f(x - 1)$. Hence

$$f(2) \;=\; 3 - f(1) \;=\; 0\,, \quad f\left(\frac{5}{2}\right) \;=\; 3 - f\left(\frac{3}{2}\right) \;=\; f\left(\frac{1}{2}\right)\,.$$

On the other hand, substituting (x, y) in (6.8) by $(2, 1/2)$, we see that

$$f\left(\frac{5}{2}\right) \;=\; f\left(\frac{1}{2}\right) + f(1) \;=\; f\left(\frac{1}{2}\right) + 3\,,$$

a contradiction.

Finally consider that $a = 1$. Then (6.9) gives $f(x) = f(x - 1) + 1$. By induction,

$$f(x+n) \;=\; f(x) + n\,, \quad \forall x \in \mathbb{R}\,, \quad \forall n \in \mathbb{Z}\,.$$

Substituting (x, y) in (6.8) by (x, n), we have then

$$n\,f(x) \;=\; f(nx)\,, \quad \forall x \in \mathbb{R}\,, \quad \forall n \in \mathbb{Z}\,.$$

Hence if $r = m/n \in \mathbb{Q}$ and $x \in \mathbb{R}$, then

$$m\,f(x) \;=\; f(mx) \;=\; f(n \cdot rx) \;=\; n\,f(rx)\,,$$

i.e. $f(rx) = r\,f(x)$. Now substituting (x, y) in (6.8) by (x, r) we get

$$f(x + r) = f(x) + r\,, \quad \forall x \in \mathbb{R}\,, \quad \forall r \in \mathbb{Q}\,.$$

Moreover, letting $x = y$ in (6.8) we get $f(x)^2 = f(x^2)$. Thus $f(x) \geq 0$ if $x \geq 0$. Since $f(-x) = -f(x)$, $f(x) \leq 0$ if $x \leq 0$. Now for any $x \in \mathbb{R}$, let $\{r_n\}$ and $\{s_n\}$ be two sequences of rational numbers increasing and decreasing to x, respectively. Then

$$r_n \leq f(x - r_n) + r_n = f(x) = f(x - s_n) + s_n \leq s_n\,,$$

which forces $f(x) = x$ by the squeeze theorem. □

Chapter 7

Equations for Trigonometric Functions

7.1 Characterization of Sine and Cosine

The reader, of course, knows several different definitions of the trigonometric functions $\sin x$ and $\cos x$:

- Using the sides of a right triangle $\triangle ABC$ (where $\hat{A} = 90°$) (see Figure 7.1) the sine and cosine of an angle are defined by:

$$\sin \hat{B} = \frac{b}{a}, \quad \cos \hat{B} = \frac{c}{a}.$$

- Using the trigonometric circle (see Figure 7.1) the sine and cosine of an angle θ are defined by:

$$\sin \theta = \overline{\mathrm{OB}}, \quad \cos \theta = \overline{\mathrm{OA}}.$$

- Using differential equations, the sine and cosine can be defined as the two linearly independent functions that solve the equation

$$y''(x) + y(x) = 0.$$

In Section 2.1, we also defined the analytic sine and analytic cosine as solutions of the equation (2.3)

$$C(x - y) = C(x)C(y) + S(x)S(y), \quad \forall x, y \in \mathbb{R}, \tag{7.1}$$

and some additional properties expecting $S(x)$ and $C(x)$ to be our well known sine and cosine functions. Indeed, in the Problem 2.14 we proved that these functions, if they exist, satisfy the exact same relations that our well known sine and cosine functions do.

In this section we return to the above functional equation and, following the work of H.E. Vaughan [62], we establish the conditions under which this equation uniquely characterizes the sine and cosine.

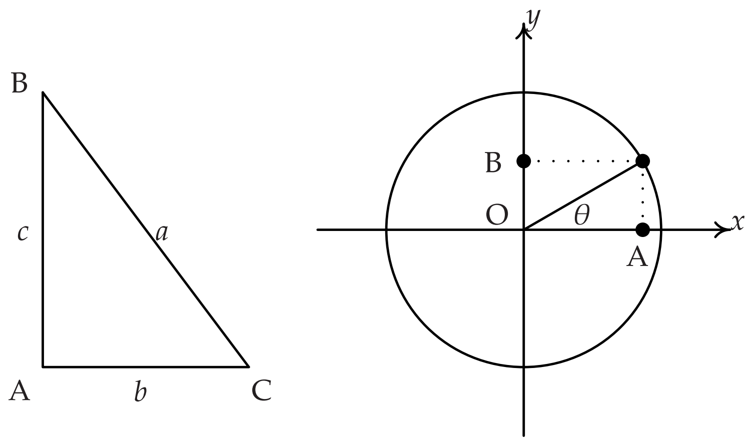

Figure 7.1: A right triangle (left) and the trigonometric circle (right) that are often used to define the trigonometric functions.

Setting $y = x$ in (7.1), we get

$$C(0) \;=\; C(x)^2 + S(x)^2 \,. \tag{7.2}$$

In particular, for $x = 0$

$$C(0) \;=\; C(0)^2 + S(0)^2 \;\Rightarrow\; C(0)\,(1 - C(0)) \;=\; S(0)^2 \,. \tag{7.3}$$

Since the right hand side is positive, the quadratic expression in the left hand side must be positive and thus

$$0 \le C(0) \le 1 \,. \tag{7.4}$$

We can now prove our first theorem:

Theorem 7.1. *If* $C(0) = 0$*, equation (7.1) has a unique solution and it is trivial:*

$$C(x) \equiv 0 \,, \quad S(x) \equiv 0 \,.$$

Proof. If $C(0) = 0$, equation (7.2) can be satisfied only if $C(x)^2 \equiv 0$ and $S(x)^2 \equiv 0$ which implies $C(x) \equiv 0$ and $S(x) \equiv 0$. □

Setting $y = 0$ in (7.1), we find

$$C(x)\,(1 - C(0)) \;=\; S(x)S(0) \,. \tag{7.5}$$

Let's square this expression and use the previous results as follows:

$$\begin{aligned}
& C(x)^2\,(1 - C(0))^2 \;=\; S(x)^2\,S(0)^2 \\
\overset{(7.3)}{\Rightarrow}\quad & C(x)^2\,(1 - C(0))^2 \;=\; S(x)^2\,C(0)(1 - C(0)) \\
\Rightarrow\quad & (1 - C(0))\,\Big[(1 - C(0))C(x)^2 - C(0)S(x)^2\Big] \;=\; 0 \\
\overset{(7.2)}{\Rightarrow}\quad & (1 - C(0))\,\Big[(1 - C(0))C(x)^2 - C(0)(C(0) - C(x)^2)\Big] \;=\; 0 \,,
\end{aligned}$$

or

$$(1 - C(0))\left[C(x)^2 - C(0)^2\right] = 0\,. \tag{7.6}$$

We have now prepared ourselves for our second theorem:

Theorem 7.2. *If* $0 < C(0) < 1$, *equation (7.1) has two solutions*

$$C(x) = c\,, \quad S(x) = b\,,$$

or

$$C(x) = c\,, \quad S(x) = -b\,,$$

where $0 < c < 1$ *and* $b = \sqrt{c - c^2}$.

Proof. If $0 < C(0) < 1$, equation (7.6) requires that $C(x) = \pm C(0)$. Consider the set

$$G = \{x \mid C(x) = C(0) \text{ or } C(x) = -C(0)\}\,.$$

Thus defined this set includes all points x in the domain of $C(x)$, that is $G = \mathbb{R}$. Endowed with the operation of addition, G is a group. Now consider the subset

$$H = \{x \mid C(x) = C(0)\}\,.$$

H is either the whole set G (if for all points $x \in \mathbb{R}$, $C(x) = C(0)$) or a proper subset of G (if there are points in $\mathbb{R}$ for which $C(x) = -C(0)$.) Let's assume that the second is true. According to a theorem of algebra, H will be a subgroup of G if, given $x, y \in H$, also $x + (-y) \in H$. If $x, y \in H$, then $C(x) = C(y) = C(0)$. Then, from equation (7.1) we see that $C(x - y) = C(0)^2 + S(0)^2$ or, using (7.3), $C(x - y) = C(0)$. So, $x - y \in H$ and, indeed, H is a subgroup. Using H, we can define an equivalence relation in G: any $a, b \in G$ are equivalent, $a \sim b$, if $a - b \in H$. This partitions G in two sets, H and $H' = \{x \mid C(x) = -C(0)\}$. Formally, one says that the index of H in G is 2. However, it is known in algebra that $\mathbb{R}$ has no such subgroup. H must thus be the whole G and

$$C(x) = C(0)\,, \quad \forall x \in \mathbb{R}\,.$$

Let $c = C(0)$. Then, from (7.2), we find that $S(x) = \pm b$, where $b = \sqrt{c - c^2}$. □

We are thus left with the case

$$C(0) = 1\,. \tag{7.7}$$

Theorem 7.3. *When* $C(0) = 1$, *the functions* $C(x)$ *and* $S(x)$ *satisfy all known relations of cosine and sine respectively.*

Proof. The condition $C(0) = 1$ now implies that

$$C(x)^2 + S(x)^2 = 1\,, \tag{7.8}$$

and

$$S(0) = 0\,. \tag{7.9}$$

Then setting $x = 0$ in the defining equation (7.1), we find that $C(x)$ is an even function

$$C(y) = C(-y)\,. \tag{7.10}$$

Setting $y = -x$ in (7.1), we find:

$$C(2x) = C(x)^2 + S(x)S(-x)\,. \tag{7.11}$$

Then, setting $y = 2x$ in the same equation and using the previous result, we get

$$\begin{aligned} C(x) &= C(x)C(2x) + S(x)S(2x) \\ &\overset{(7.11)}{=} C(x)\left(C(x)^2 + S(x)S(-x)\right) + S(x)S(2x) \\ &= C(x)^3 + C(x)S(x)S(-x) + S(x)S(2x)\,, \end{aligned}$$

which we solve for $S(x)S(2x)$:

$$\begin{aligned} S(x)S(2x) &= C(x)\left[(1 - C(x)^2) - S(x)S(-x)\right] \\ &\overset{(7.8)}{=} C(x)\left[S(x)^2 - S(x)S(-x)\right] \\ &= C(x)S(x)\,[S(x) - S(-x)]\,. \end{aligned}$$

If $S(x) \neq 0$, we conclude from this equation that

$$S(2x) = C(x)\,[S(x) - S(-x)]\,. \tag{7.12}$$

We will show that this equation is also true when $S(x) = 0$ and thus true for any x. From equation (7.8), we see that

$$S(x)^2 = S(-x)^2\,.$$

If $S(x) = 0$, then $S(-x) = 0$ too. Also $C(x)^2 = 1$ (from (7.8)) and consequently $C(2x) = 1$ (from (7.11)) and $S(2x) = 0$ (from (7.8)). Therefore, when $S(x) = 0$ equation (7.12) is the trivial identity $0 = 0$.

From the last equation, we can see that for any $x/2$

$$\text{either } S(x/2) = S(-x/2) \quad \text{or } S(x/2) = -S(-x/2)\,.$$

If $S(x/2) = S(-x/2)$, then $S(x) = 0$ (from (7.12)) and also $S(-x) = 0$. If $S(x/2) = -S(-x/2)$, then, with the help of equation (7.12),

$$S(x) = 2\,C(x/2)\,S(x/2)\,. \tag{7.13}$$

Substituting $-x$ for x in this equation we find

$$S(-x) \;=\; -\,2\,C(x/2)\,S(-x/2)\;,$$

which establishes that $S(x) = -S(-x)$ for $S(x) \neq 0$. Taking into account all the previous discussion we conclude that the function $S(x)$ is odd

$$S(x) \;=\; -\,S(-x)\;, \tag{7.14}$$

for all x and that equation (7.13) is also true for all x.

Then, setting $-y$ for y in (7.1) we can immediately find

$$C(x+y) \;=\; C(x)C(y) - S(x)S(y)\;. \tag{7.15}$$

Setting $x+y$ for y in (7.1)

$$\begin{aligned} C(-y) &= C(x)C(x+y) + S(x)S(x+y) \\ &\overset{(7.15)}{=} C(x)\left[C(x)C(y) - S(x)S(y)\right] + S(x)S(x+y)\;. \end{aligned}$$

Solving for $S(x)S(x+y)$:

$$\begin{aligned} S(x)S(x+y) &= C(y)(1 - C(x)^2) + C(x)S(x)S(y) \\ &= C(y)S(x)^2 + C(x)S(x)S(y) \\ &= S(x)\left[S(x)C(y) + S(y)C(x)\right]\;, \end{aligned}$$

and, if $S(x) \neq 0$,

$$S(x+y) \;=\; S(x)C(y) + S(y)C(x)\;. \tag{7.16}$$

Obviously this equation is also true if $S(y) \neq 0$ (by renaming the variables). If $S(x) = S(y) = 0$ then $C(x), C(y) = \pm 1 \Rightarrow C(x+y) = \pm 1 \Rightarrow S(x+y) = 0$ and the above equation is also true.

Setting $-y$ for y in (7.16), we also find

$$S(x-y) \;=\; S(x)C(y) - S(y)C(x)\;. \tag{7.17}$$

We can now derive any other formula that duplicates any known trigonometric formula. For example:

$$\begin{aligned} C(2x) &= C(x)^2 - S(x)^2 \;=\; 2C(x)^2 - 1 \;=\; 1 - 2S(x)^2\;, \\ S(2x) &= 2S(x)C(x)\;, \\ S(3x) &= S(x)\left[4C(x)^2 - 1\right] \;=\; S(x)\left[3 - 4S(x)^2\right]\;, \\ C(x) - C(y) &= 2S\left(\frac{x+y}{2}\right)S\left(\frac{y-x}{2}\right)\;, \\ C(x) + C(y) &= 2C\left(\frac{x+y}{2}\right)C\left(\frac{x-y}{2}\right)\;, \\ S(x) - S(y) &= 2S\left(\frac{x-y}{2}\right)C\left(\frac{x+y}{2}\right)\;, \\ S(x) + S(y) &= 2C\left(\frac{x-y}{2}\right)S\left(\frac{x+y}{2}\right)\;, \end{aligned}$$

and so on. □

Although we have shown that $S(x)$ and $C(x)$ satisfy all usual relations of trigonometry, we have not yet shown that they are actually the familiar sine and cosine functions. In fact, equation (7.1) can have solutions that are discontinuous[1] at every point.

Theorem 7.4. *If* $\lim_{x\to 0+} C(x)$ *exists, then* $\lim_{x\to 0} S(x) = 0$.

Proof. Let $\lim_{x\to 0+} C(x) = \ell$. Then from $C(2x) = 2C(x)^2 - 1$ it must be that $\ell = 2\ell^2 - 1$ or $\ell = 1, -1/2$.

Since $|S(x)| \leq 1$, the $\lim_{x\to 0^+} S(x)$ cannot diverge. Then, if $\ell = -1/2$, from the relation $S(3x) = S(x)[4C(x)^2 - 1]$, as $x \to 0^+$, the right hand side converges to 0 and thus $\lim_{x\to 0^+} S(x) = 0$. However, from $C(x)^2 + S(x)^2 = 1$, if $\ell = -1/2$, we see that $\lim_{x\to 0^+} S(x)^2 = 3/4$ which contradicts the previous result. Therefore, ℓ must be 1. Since $C(x)$ is even the limit from the other side exist and it is equal:

$$\lim_{x\to 0} C(x) \;=\; 1\,.$$

Similarly, since $S(x)$ is odd

$$\lim_{x\to 0} S(x) \;=\; 0\,.$$

Given that $C(0) = 1$ and $S(0) = 0$, the above relations imply that $C(x)$ and $S(x)$ are continuous at $x = 0$. □

Notice that one could have assumed that $\lim_{x\to 0^+} S(x)$ exists to reach the same conclusions.

Theorem 7.5. *If $S(x)$ (or $C(x)$) is continuous at a point (in particular $x = 0$), then both functions are continuous at all points.*

Proof. If $S(x)$ is continuous at 0, as $y \to x$ the limit $\lim S(\frac{x-y}{2})$ exists and it is equal to $S(0) = 0$. Then, in the same limit, the right hand side of the equation

$$C(x) - C(y) \;=\; -2S\left(\frac{x+y}{2}\right)S\left(\frac{x-y}{2}\right) \tag{7.18}$$

converges to 0 and thus

$$\lim_{y\to x} C(y) \;=\; C(x)\,.$$

Using the rest of the identities that change a sum to a product, we can prove the remaining statements. □

Theorem 7.6. *$S(x)$ is continuous if and only if there exists a $\lambda > 0$ such that $S(x)$ does not change sign in the interval $(0, \lambda)$.*

[1]See Chapter 11 to fully understand this statement. In fact, you may want to read the remaining section only after you have read Chapter 11 and, perhaps, Chapter 12.

Proof. Necessary: Let $S(x)$ be continuous. Imagine that we cannot find an interval that $S(x)$ maintains its sign. Then, for each possible interval there must be a zero of $S(x)$. We consider the set $K = \{x \mid f(x) = 0\}$ of zeros. For any $x \in \mathbb{R}$ we can find a sequence $\{x_n\}$ in K such that $x_n \to x$. By continuity $S(x) = S(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} S(x_n) = \lim_{n\to\infty} 0 = 0$. That is, $S(x)$ is identically zero. However, this cannot be true since we are searching for non-trivial solutions. Therefore, there must be an interval in which $S(x)$ maintains its sign.

Sufficient: Let $S(x)$ maintain its sign in the interval $(0, \lambda)$. Consider two arbitrary points x, y in this interval such that $y < x$. Then $\frac{x-y}{2}, \frac{x+y}{2} \in (0, \lambda)$ and the right hand side of equation (7.18) is negative. The left hand side must also be negative or $C(x) < C(y)$, implying that $C(x)$ is a decreasing function in $(0, \lambda)$. The limit $\lim_{x\to 0^+} C(x)$ thus exists and, according to the theorems proved above, $S(x)$ is continuous. □

Theorem 7.7. *If $C(x)$ is continuous, then $S(x)$ is differentiable and*

$$S'(x) \;=\; S'(0)\, C(x)\,. \tag{7.19}$$

Proof. Since $C(x)$ is continuous, $S(x)$ is continuous too, and there is an interval $(0, \lambda)$ such that $S(x)$ maintains its sign. Let $x \in (0, \lambda)$ and set $\varepsilon = x/n$ (that is, partition the interval $(0, x)$ into n smaller intervals). Then

$$\begin{aligned} S\left(\frac{\varepsilon}{2}\right) \sum_{k=1}^{n} C(k\varepsilon) &= \sum_{k=1}^{n} C\left((2k-1)\frac{\varepsilon}{2} + \frac{\varepsilon}{2}\right) S\left(\frac{\varepsilon}{2}\right) \\ &= \frac{1}{2} \sum_{k=1}^{n} \left[S\left((2k+1)\frac{\varepsilon}{2}\right) - S\left((2k-1)\frac{\varepsilon}{2}\right)\right] \\ &= \frac{1}{2}\left[S\left(x + \frac{\varepsilon}{2}\right) - S\left(\frac{\varepsilon}{2}\right)\right], \end{aligned}$$

where in the second equality we used the identity

$$S(a+b) - S(a-b) \;=\; 2\, C(a)\, S(b)\,.$$

Now we rewrite the above sum in the form

$$\sum_{k=1}^{n} C(k\varepsilon)\, \varepsilon \;=\; \left[S\left(x + \frac{\varepsilon}{2}\right) - S\left(\frac{\varepsilon}{2}\right)\right] \frac{\varepsilon/2}{S(\varepsilon/2)}\,.$$

In the limit $n \to \infty$, $\varepsilon \to 0$

$$\int_0^x C(u)\, du \;=\; S(x) \lim_{x\to 0} \frac{x}{S(x)}\,.$$

Since $C(0) = 1$, in a neighborhood $(0, x)$ of 0 with x sufficiently small, $C(u) > 0$ and $S(x) \neq 0$. The integral is then non-zero and hence the limit in the right hand side exists and it is non-zero. In particular,

$$\lim_{x\to 0} \frac{S(x)}{x} \;=\; \lim_{x\to 0} \frac{S(x) - S(0)}{x - 0} \;=\; S'(0)\,.$$

Therefore

$$\int_0^x C(u)du = \frac{S(x)}{S'(0)} .$$

The derivative of the integral in the left hand side exists and therefore the derivative of the right hand side exists too, resulting to equation (7.19). □

Similarly, we can prove:

Theorem 7.8. *If $S(x)$ is continuous, then $C(x)$ is differentiable and*

$$C'(x) = -S'(0)\,S(x) . \tag{7.20}$$

For $x = 0$, equations (7.19) and (7.20) give:

$$S'(0) = S'(0)C(0) = S'(0) , \quad C'(0) = -S'(0)S(0) = 0 .$$

That is, the derivative $S'(0)$ remains undefined and $C'(0) = 0$. Let's set $a = S'(0)$.

We are now ready to identify the functions $S(x)$ and $C(x)$ with known functions. We can do it in at least two ways. The first is a little simpler but it assumes some knowledge of the theory of differential equations. Both methods rely on results from calculus. The right hand sides of equations (7.19) and (7.20) are differentiable. Therefore $C(x)$ and $S(x)$ are twice-differentiable:

$$\begin{aligned} S''(x) &= a\,C'(x) = -a^2\,S(x) , \\ C''(x) &= -a\,S'(x) = -a^2\,C(x) . \end{aligned}$$

Therefore the two functions $S(x)$, $C(x)$ satisfy the differential equation

$$y''(x) + ay(x) = 0 ,$$

which is solved by

$$C(x) = \cos(ax) , \quad S(x) = \sin(ax) .$$

The second way to identify the functions $S(x)$, $C(x)$ is through their Taylor series. Inductively, we can easily see that $C(x)$ and $S(x)$ have derivatives of any order. In particular

$$\begin{aligned} S^{(2n)}(x) &= (-1)^n\,a^{2n}\,S(x) , \\ S^{(2n+1)}(x) &= (-1)^n\,a^{2n+1}\,C(x) . \end{aligned}$$

Then

$$S^{(2n)}(0) = 0 , \quad S^{(2n+1)} = (-1)^n a^{2n+1} ,$$

and we can write the formal Taylor series for $S(x)$:

$$S(x) = \sum_{k=0}^{\infty} \frac{1}{k!}\,S^{(k)}(0)\,x^k = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\,(ax)^{2n+1} .$$

This series converges and it has the limit $\sin(ax)$, so $S(x) = \sin(ax)$. Similarly, we can show that

$$C(x) \;=\; \sum_{k=0}^{\infty} \frac{1}{k!}\, C^{(k)}(0)\, x^k \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\, (ax)^{2n} \;,$$

which also converges, and has the limit $C(x) = \cos(ax)$.

Taking into account everything that has been discussed in this section, we can now state the following theorem:

Theorem 7.9. *The functional equation (7.1) with the conditions*
(a) $C(0) = 1$;
(b) $\lim_{x\to 0^+} C(x)$ *exists;*
(c) $\lim_{x\to 0^+} \frac{S(x)}{x} = 1$,
has a unique solution

$$S(x) \;=\; \sin x \;, \quad C(x) \;=\; \cos x \;.$$

Returning to the analytic sine and cosine, we can now see that they result in a unique solution of (7.1): $S(x) = \sin x$, $C(x) = \cos x$. However the conditions imposed are slightly stronger that the minimal conditions required.

7.2 D'Alembert-Poisson I Equation

Problem 7.1. *Find all continuous functions* $f : \mathbb{R} \to \mathbb{R}$ *satisfying*

$$f(x+y) + f(x-y) \;=\; 2\, f(x)\, f(y) \;, \quad \forall x, y \in \mathbb{R} \;. \tag{7.21}$$

Comment. Equation (7.21) is sometimes referred to as the **D'Alembert functional equation** [7], sometimes as the **Poisson functional equation** [33], and occasionally as the **cosine functional equation** [50]. I will call it the **D'Alembert-Poisson I functional equation**.

Solution. The equation (7.21) may be solved using the $\mathbb{NQR}$-method as described in chapter 6. However, it admits trigonometric solutions, such as $\cos(bx)$ and if we follow the $\mathbb{NQR}$-method verbatim we shall face difficulties that arise from the addition properties of $\cos(mx/n)$, for integer m, n. For this reason, Cauchy himself presented a modification of his $\mathbb{NQR}$-method that simplified the problem considerably. This is the method we follow here.

Setting $y = 0$ in (7.21), we find

$$2f(x) \;=\; 2f(x)f(0) \;.$$

Therefore either $f(x) \equiv 0$ or $f(0) = 1$. We shall investigate the second case. Since $f(x)$ is continuous and $f(0) = 1$, there exists $\varepsilon > 0$ such that $f(x) > 0$, $\forall x \in [-\varepsilon, +\varepsilon]$.

Setting now $x = 0$ in (7.21), we find

$$f(y) + f(-y) \; = \; 2f(0)f(y) \; \Leftrightarrow \; f(-y) \; = \; f(y) \, .$$

Therefore the function f is even and we only need to find its values for $x > 0$. Initially, we shall find the value of f for the points $\varepsilon, \frac{\varepsilon}{2}, \frac{\varepsilon}{4}, \cdots \in [0, +\varepsilon]$. Towards this goal, let $y = x$ in (7.21):

$$f(x)^2 \; = \; \frac{1 + f(2x)}{2} \, . \tag{7.22}$$

At $x = \varepsilon$, we know that $f(\varepsilon) > 0$. There are two possibilities. Either $f(\varepsilon) \leq 1$ or $f(\varepsilon) > 1$. We shall examine the first case as the second case is done similarly. We define a number $\theta_0 \in [0, \pi/2]$ such that $f(\varepsilon) = \cos\theta_0$.

Setting $x = \varepsilon/2$ in (7.22) we find

$$f(\varepsilon/2)^2 \; = \; \frac{1 + f(\varepsilon)}{2} \; = \; \frac{1 + \cos\theta_0}{2} \; = \; \cos^2\frac{\theta_0}{2} \; \Leftrightarrow \; f(\varepsilon/2) \; = \; \cos\frac{\theta_0}{2} \, ,$$

since both are positive. Then, setting $x = \varepsilon/4$ in (7.22) we find

$$f(\varepsilon/4)^2 \; = \; \frac{1 + f(\varepsilon/2)}{2} \; = \; \frac{1 + \cos(\theta_0/2)}{2} \; = \; \cos^2\frac{\theta_0}{4} \; \Leftrightarrow \; f(\varepsilon/4) \; = \; \cos\frac{\theta_0}{4} \, .$$

Continuing like this we have

$$f\left(\frac{\varepsilon}{2^n}\right) \; = \; \cos\frac{\theta_0}{2^n} \, , \quad \forall n \in \mathbb{N}^* \, .$$

Now we shall evaluate the function f at the points $m\varepsilon/2^n$, $m, n \in \mathbb{N}^*$. This will take us out of the interval $[0, \varepsilon]$. To implement this, let's set $x = ny$ in (7.21) to find

$$f\,((n+1)y) \; = \; 2\,f(ny)\,f(y) - f((n-1)y) \, . \tag{7.23}$$

In this equation we set $n = 2$, $y = \varepsilon/2^n$:

$$\begin{aligned} f\left(\frac{3\varepsilon}{2^n}\right) \; &= \; 2f\left(\frac{\varepsilon}{2^n}\right) f\left(\frac{3\varepsilon}{2^{n-1}}\right) - f\left(\frac{\varepsilon}{2^n}\right) \\ &= \; 2\cos\frac{\theta_0}{2^n}\cos\frac{\theta_0}{2^{n-1}} - \cos\frac{\theta_0}{2^n} \\ &= \; \cos\frac{3\theta_0}{2^n} \, , \end{aligned}$$

since

$$\begin{aligned} \cos(3x) \; &= \; \cos(2x + x) \; = \; \cos(2x)\cos x - \sin(2x)\sin x \; = \; \cos(2x)\cos x - 2\sin^2 x\cos x \\ &= \; \cos(2x)\cos x - 2\frac{1 - \cos(2x)}{2}\cos x \; = \; 2\cos(2x)\cos x - \cos x \, . \end{aligned}$$

Inductively

$$f\left(\frac{m\varepsilon}{2^n}\right) = \cos\frac{m\theta_0}{2^n}\,. \tag{7.24}$$

Let $x \in \mathbb{R}_+$. We take a sequence $\frac{m\varepsilon}{2^n}$, $m, n \in \mathbb{N}^*$, such that $\lim\limits_{\substack{m\to\infty\\ n\to\infty}} \frac{m\varepsilon}{2^n} = x$. Then $\lim\limits_{\substack{m\to\infty\\ n\to\infty}} \frac{m}{2^n} = \frac{x}{\varepsilon}$ and since f is continuous,

$$f(x) = \cos\left(\frac{\theta_0}{\varepsilon}x\right), \quad x \in \mathbb{R}_+.$$

We define $a = \theta_0/\varepsilon$ and, because f is even,

$$f(x) = \cos(ax)\,, \quad x \in \mathbb{R}\,.$$

The case $f(\varepsilon) > 1$ is studied similarly. We define a number ξ_0 such that $f(\varepsilon) = \cosh\xi_0$. Repeating the steps presented above,

$$f(x) = \cosh(ax)\,, \quad x \in \mathbb{R}\,.$$ □

7.3 D'Alembert-Poisson II Equation

Problem: Find all continuous functions $f : \mathbb{R} \to \mathbb{R}$ which satisfying

$$f(x+y)\,f(x-y) = f(x)^2 - f(y)^2\,, \quad \forall x, y \in \mathbb{R}\,. \tag{7.25}$$

Comment. Occasionally as the this equation is called the **sine functional equation** [50]. To match my terminology of equation (7.21), I will refer to equation (7.25) as the **D'Alembert-Poisson II functional equation**.

Comment. Equation (7.25) may be written equivalently

$$f(x)\,f(y) = f\left(\frac{x+y}{2}\right)^2 - f\left(\frac{x-y}{2}\right)^2\,, \quad \forall x, y \in \mathbb{R}\,. \tag{7.26}$$

Solution. An obvious solution is $f(x) \equiv 0$. We shall search for solutions that do not vanish identically. Therefore, there exists $x_0 \in \mathbb{R}$ such that $f(x_0) \neq 0$. We can thus define

$$g(x) = \frac{f(x+x_0) - f(x-x_0)}{2f(x_0)}\,. \tag{7.27}$$

Defined like this, it seems that g depends on the point x_0 selected. However, this is not true. The function g is independent of x_0. To see this, let $\tilde{x}_0$ be another point for which $f(\tilde{x}_0) \neq 0$. We define

$$\tilde{g}(x) = \frac{f(x+\tilde{x}_0) - f(x-\tilde{x}_0)}{2f(\tilde{x}_0)}\,.$$

We will show that $g(x) \equiv \tilde{g}(x)$. Indeed, multiplying and dividing the right hand side of the equation that defines $\tilde{g}(x)$ and then using equation (7.26) twice, we find:

$$\begin{aligned}
\tilde{g}(x) &= \frac{f(x+\tilde{x}_0)f(x_0) - f(x-\tilde{x}_0)f(x_0)}{2f(\tilde{x}_0)f(x_0)} \\
&\overset{(7.26)}{=} \frac{f(\frac{x+\tilde{x}_0+x_0}{2})^2 - f(\frac{x+\tilde{x}_0-x_0}{2})^2 - f(\frac{x-\tilde{x}_0+x_0}{2})^2 + f(\frac{x-\tilde{x}_0-x_0}{2})^2}{2f(\tilde{x}_0)f(x_0)} \\
&= \frac{\left[f(\frac{x+x_0+\tilde{x}_0}{2})^2 - f(\frac{x-x_0+\tilde{x}_0}{2})^2\right] - \left[f(\frac{x-x_0+\tilde{x}_0}{2})^2 - f(\frac{x-x_0-\tilde{x}_0}{2})^2\right]}{2f(\tilde{x}_0)f(x_0)} \\
&\overset{(7.26)}{=} \frac{f(x+x_0)f(\tilde{x}_0) - f(x-x_0)f(\tilde{x}_0)}{2f(\tilde{x}_0)f(x_0)} \\
&= \frac{f(x+x_0) - f(x-x_0)}{2f(x_0)} = g(x)\,.
\end{aligned}$$

The function $g(x)$ satisfies (7.21). To see this,

$$\begin{aligned}
2g(x)g(y) &= 2\left(\frac{f(x+x_0) - f(x-x_0)}{2f(x_0)}\right)\left(\frac{f(y+x_0) - f(y-x_0)}{2f(x_0)}\right) \\
&= \frac{1}{2f(x_0)^2}\,[f(x+x_0)f(y+x_0) - f(x+x_0)f(y-x_0) \\
&\qquad -f(x-x_0)f(y+x_0) + f(x-x_0)f(y-x_0)]\,.
\end{aligned}$$

Let

$$u = \frac{x+y}{2}, \quad v = \frac{x-y}{2} \;\Rightarrow\; x = u+v, \quad y = u-v\,.$$

Then

$$\begin{aligned}
2g(x)g(y) &= \frac{1}{2f(x_0)^2}\,[f(u+v+x_0)f(u-v+x_0) - f(u+v+x_0)f(u-v-x_0) \\
&\qquad -f(u+v-x_0)f(u-v+x_0) + f(u+v-x_0)f(u-v-x_0)] \\
&= \frac{1}{2f(x_0)^2}\,[f((u+x_0)+v)f((u+x_0)-v) - f(u+(v+x_0))f(u-(v+x_0)) \\
&\qquad -f(u+(v-x_0))f(u-(v-x_0)) + f((u-x_0)+v)f((u-x_0)-v)]
\end{aligned}$$

We can use (7.25) to rewrite the right hand side of the last equation:

$$2g(x)g(y) = \frac{f(u+x_0)^2 - f(v)^2 - f(u)^2 + f(v+x_0)^2 - f(u)^2 + f(v-x_0)^2 + f(u-x_0)^2 - f(v)^2}{2f(x_0)^2}\,.$$

Now, we re-arrange the terms on the right hand side and use (7.25) once more:

$$\begin{aligned}
2g(x)g(y) &= \frac{f(u+x_0)^2 - f(u)^2 - f(v)^2 + f(v+x_0)^2 - f(v)^2 + f(v-x_0)^2 + f(u-x_0)^2 - f(u)^2}{2f(x_0)^2} \\
&= \frac{f(2u+x_0)\,f(x_0) - f(2v+x_0)\,f(x_0) - f(2v-x_0)\,f(x_0) + f(2u-x_0)\,f(x_0)}{2f(x_0)^2} \\
&= \frac{f(2u+x_0) - f(2v+x_0) - f(2v-x_0) + f(2u-x_0)}{2f(x_0)} \\
&= \frac{f(2u+x_0) - f(2u-x_0)}{2f(x_0)} + \frac{f(2v+x_0) - f(2v-x_0)}{2f(x_0)} \\
&= g(2u) + g(2v)\,,
\end{aligned}$$

or

$$2g(x)g(y) \;=\; g(x+y) + g(x-y)\,.$$

Since g is continuous,

$$g(x) \;\equiv\; 0\,, \quad g(x) \;=\; \cos(bx)\,, \quad g(x) \;=\; \cosh(bx)\,.$$

Knowing $g(x)$, we can invert equation (7.27):

$$f(x+y) - f(x-y) \;=\; 2\,f(y)\,g(x)\,,$$

for all $x \in \mathbb{R}$ and for all $y \in \mathbb{R}$ such that $f(y) \neq 0$.

For $g(x) \;\equiv\; 0$,

$$f(x+y) - f(x-y) \;=\; 0 \;\Rightarrow\; f(x+2y) \;=\; f(x) \;\Rightarrow\; f(2y) \;=\; f(0)\,.$$

From (7.25) we can see that $f(0) = 0$ and therefore $f(x) \equiv 0$. This solution is not consistent with our assumptions.

For $g(x) \;=\; \cos(bx)$,

$$f(x+y) - f(x-y) \;=\; 2\,f(y)\,\cos(bx)\,.$$

For $x = 0$, we find

$$f(y) - f(-y) \;=\; 2\,f(y) \;\Rightarrow\; f(-y) \;=\; -\,f(y)\,.$$

Therefore,

$$f(y+x) + f(y-x) \;=\; 2\,f(y)\,\cos(bx)\,.$$

This equation has been solved in Problem 4.6 and the solution is

$$f(x) \;=\; A\,\sin(bx) + B\,\cos(bx)\,.$$

Since f must be odd,

$$f(x) \;=\; A\,\sin(bx)$$

is the final result.

For $g(x) \;=\; \cosh(bx)$, we find similarly

$$f(x) \;=\; A\,\sinh(bx)\,. \qquad \square$$

7.4 Solved Problems

Problem 7.2. *Consider the function $C : \mathbb{R} \to \mathbb{R}$ satisfying the conditions:*
(A) It is continuous at all points $x \in \mathbb{R}$.
(B) It satisfies the functional equation

$$C(x+y) + C(x-y) = 2\,C(x)\,C(y) \tag{7.28}$$

(C) There is a λ which is the least non-negative root of $C(x) = 0$.
(D) $C(0) > 0$.
Prove that
(a) $C(0) > 0$ for all $x \in (0, \lambda)$.
(b) $C(0) = 1$.
(c) $C(-x) = C(x)$, for all $x \in \mathbb{R}$.
(d) $C(x + 2\lambda) = -C(x)$.
(e) $C(2x) = 2C(x)^2 - 1$ for all $x \in \mathbb{R}$.
(f) $C(x/2) = \pm\sqrt{\frac{1+C(x)}{2}}$ for all $x \in \mathbb{R}$.
(g) $|C(x)| \leq 1$ for all $x \in \mathbb{R}$.
(h) The function $C(x)$ is periodic with least positive period 4λ.

Solution. (a) Condition (D) implies that $\lambda > 0$. Since λ is the least positive root of $C(x)$, for any $x \in (0, \lambda)$, $C(x) \neq 0$. The continuity of $C(x)$ and the fact that $C(0) > 0$, imply that $C(x) > 0$ for any $x \in (0, \lambda)$.

(b) Setting $y = 0$ in the defining equation (7.28), we find

$$2C(x) = 2C(x)\,C(0) \Rightarrow C(x)\,(C(0) - 1) = 0\,.$$

This equation is true for all x and, in particular, for $x \in (0, \lambda)$. In the latter case, $C(x) \neq 0$ and thus $C(0) = 1$.

(c) Setting $x = 0$ in equation (7.28), we find

$$C(y) + C(-y) = 2C(0)\,C(y)\,.$$

Combining with the previous result, $C(0) = 1$, it immediately follows that $C(x)$ is even.

(d) In equation (7.28), we substitute $x + \lambda$ in place of x and λ for y:

$$C(x + 2\lambda) + C(x) = 2C(x + \lambda)\,C(\lambda) \Rightarrow C(x + 2\lambda) + C(x) = 0\,,$$

which is the result sought.

(e) For $y = x$, the defining equation (7.28) immediately gives $C(2x) + C(0) = 2C(x)^2$ or $C(2x) = 2C(x)^2 - 1$.

(f) Solving for $C(x)$ the previous relation

$$C(x) = \pm\sqrt{\frac{1 + C(2x)}{2}}\,.$$

The relation asked is found upon replacing x with $x/2$.

(g) Let's assume that the statement $|C(x)| \leq 1$, $\forall x$ is not true. Then there is at least a number $a \in \mathbb{R}$ such that $|C(a)| > 1$.

Let $x = \lambda + a$ and $y = \lambda - a$. Then equation (7.28) gives

$$C(2\lambda) + C(2a) \;=\; 2C(\lambda + a)\,C(\lambda - a)\,. \tag{7.29}$$

The left hand can be rewritten as

$$[2C(\lambda)^2 - 1] + [2C(a)^2 - 1] \;=\; 2[C(a)^2 - 1]\,.$$

This is a positive number. The right hand side can be rewritten as

$$2C(\lambda + a)C(2\lambda - (\lambda + a)) \;=\; -2C(\lambda + a)C(-(\lambda + a)) \;=\; -C(\lambda + a)^2\,.$$

This is a negative number and therefore equation (7.29) is impossible. Therefore we must have $|C(x) \leq 1$, $\forall x$.

(h) That the function is periodic is easy to see by a double application of the result (d):

$$C(4\lambda + x) \;=\; C(2\lambda + (2\lambda + x)) \;=\; -C(2\lambda + x) \;=\; C(x)\,.$$

It only remains to prove that 4λ is the least period. Let's assume that this is not true and therefore there is a period T such that $0 < T < 4\lambda$. Then

$$C(T) \;=\; C(0 + T) \;=\; C(0) \;=\; 1\,.$$

We now apply equation (7.28) for $x = y = T/2$:

$$C(T) + C(0) \;=\; 2\,C\left(\frac{T}{2}\right)^2 \;\Rightarrow\; C\left(\frac{T}{2}\right) \;=\; \pm 1\,.$$

If $C(T/2) = -1$, we set $x = y = T/4$ in equation (7.29) to find

$$C\left(\frac{T}{2}\right) + C(0) \;=\; 2\,C\left(\frac{T}{4}\right)^2 \;\Rightarrow\; C\left(\frac{T}{4}\right) \;=\; 0\,.$$

Since $0 < T/4 < \lambda$, this is impossible. If $C(T/2) = 1$, we set $x = \lambda + T/4$, $y = \lambda - T/4$ in equation (7.29) to find

$$C(2\lambda) + C\left(\frac{T}{2}\right) \;=\; 2\,C\left(\lambda + \frac{T}{4}\right)C\left(\lambda - \frac{T}{4}\right) \;\Rightarrow\; C\left(\lambda + \frac{T}{4}\right)C\left(\frac{\lambda - T}{4}\right) \;=\; 0\,,$$

since $C(2\lambda) = -1$. We also notice that

$$C\left(\lambda + \frac{T}{4}\right) \;=\; C\left(2\lambda - \left(\lambda - \frac{T}{4}\right)\right) \;=\; -C\left(\lambda - \frac{T}{4}\right)\,.$$

The previous equation is thus equivalent to

$$C\left(\lambda - \frac{T}{4}\right)^2 \;=\; 0 \;\Rightarrow\; C\left(\lambda - \frac{T}{4}\right) \;=\; 0\,.$$

However, since $0 < \lambda - T/4 < \lambda$, this is impossible too. We thus conclude that 4λ is the least period. $\square$

Part III

GENERALIZATIONS

Chapter 8

Pexider, Vincze & Wilson Equations

Cauchy's functional equations

$$\begin{aligned} f(x+y) &= f(x)+f(y)\,, \\ f(x+y) &= f(x)\,f(y)\,, \\ f(x\,y) &= f(x)+f(y)\,, \\ f(x\,y) &= f(x)\,f(y)\,, \end{aligned}$$

can be generalized as relations among three functions:

$$\begin{aligned} f(x+y) &= g(x)+h(y)\,, \\ f(x+y) &= g(x)\,h(y)\,, \\ f(x\,y) &= g(x)+h(y)\,, \\ f(x\,y) &= g(x)\,h(y)\,, \end{aligned}$$

Even more, one can generalize them to include an arbitrary number of functions:

$$\begin{aligned} f(x_1+x_2+\cdots+x_n) &= f_1(x_1)+f_2(x_2)+\cdots+f_n(x_n)\,, \\ f(x_1+x_2+\cdots+x_n) &= f_1(x_1)\,f_2(x_2)\,\cdots\,f_n(x_n)\,, \\ f(x_1\,x_2\,\cdots\,x_n) &= f_1(x_1)+f_2(x_2)+\cdots+f_n(x_n)\,, \\ f(x_1\,x_2\,\cdots\,x_n) &= f_1(x_1)\,f_2(x_2)\,\cdots\,f_n(x_n)\,. \end{aligned}$$

8.1 First Pexider Equation

In order to understand fully Pexider's equations, let's first understand the original equation containing three functions and then we will solve the most general version of it.

Problem 8.1. *Let $f, g, h : \mathbb{R} \to \mathbb{R}$ be three continuous functions that satisfy the functional relation*

$$f(x+y) = g(x)+h(y)\,, \qquad \forall x, y\,. \tag{8.1}$$

Find all such functions f, g, h.

Solution. Setting $x = y = 0$ in the defining equation, we see that

$$f(0) = g(0) + h(0) . \tag{8.2}$$

Also, if $y = 0$ or $x = 0$,

$$\begin{aligned} f(x) &= g(x) + h(0) , && (8.3) \\ f(y) &= g(0) + h(y) . && (8.4) \end{aligned}$$

Adding the last two equations and using the defining equation and (8.2), we find

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

or

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

That is, the function $f(x) - f(0)$ satisfies the linear Cauchy equation (5.1) and therefore f is given by

$$f(x) = c\,x + a + b , \quad c \in \mathbb{R} .$$

In the above equation we set $g(0) = a$, $h(0) = b$ and thus $f(0) = a + b$. From (8.3) and (8.4), we find then

$$g(x) = c\,x + a , \qquad h(x) = c\,x + b . \qquad \square$$

Having understood the simple case of three different functions we can now solve the problem in full generality.

Problem 8.2. *Let $f, f_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, \ldots, n$, be $n + 1$ continuous functions that satisfy the functional relation*

$$f(x_1 + x_2 + \cdots + x_n) = f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n) , \tag{8.5}$$

$\forall x_1, x_2, \ldots, x_n$. *Find all such functions f, f_i.*

Comment. Equation (8.5) (and thus the simpler case (8.1) too) is known as the **first Pexider functional equation** or the **linear Pexider functional equation**.

Solution. Setting $x_1 = x_2 = \cdots = x_n = 0$ in the defining equation, we see that

$$f(0) = f_1(0) + f_2(0) + \cdots + f_n(0) . \tag{8.6}$$

Also, we set successively $x_1 = x_2 = \cdots = x_{i-1} = x_{i+1} = \cdots = x_n = 0$, for $i = 1, 2, \ldots, n$:

$$\begin{aligned} f(x_1) &= f_1(x_1) + f_2(0) + f_3(0) + \cdots + f_{n-1}(0) + f_n(0) , \\ f(x_2) &= f_1(0) + f_2(x_2) + f_3(0) + \cdots + f_{n-1}(0) + f_n(0) , \\ \cdots &\quad \cdots \\ f(x_n) &= f_1(0) + f_2(0) + f_3(0) + \cdots + f_{n-1}(0) + f_n(x_n) . \end{aligned}$$

Adding the last n equations and using the defining equation and (8.6), we find

$$f(x_1) + f(x_2) + \cdots + f(x_n) \;=\; f(x_1 + x_2 + \cdots + x_n) + (n-1)\, f(0)\ ,$$

or

$$[f(x_1) - f(0)] + \cdots + [f(x_n) - f(0)] \;=\; [f(x_1 + x_2 + \cdots + x_n) - f(0)]\ .$$

That is, the function $f(x) - f(0)$ satisfies equation (5.3) which is equivalent to the linear Cauchy equation (5.1) and therefore f is given by

$$f(x) \;=\; c\,x + \sum_{i=1}^{n} a_i\ , \quad c \in \mathbb{R}\ ,$$

where we set $f_i(0) = a_i$, $\forall i$. Then

$$f_i(x) \;=\; c\,x + a_i\ , \quad \forall i = 1, 2, \ldots, n\ . \qquad \square$$

8.2 Second Pexider Equation

Problem 8.3. *Let* $f, f_i : \mathbb{R} \to \mathbb{R}$*,* $i = 1, 2, \ldots, n$*, be* $n+1$ *continuous functions that satisfy the functional equation*

$$f(x_1 + x_2 + \cdots + x_n) \;=\; f_1(x_1)\, f_2(x_2) \ldots f_n(x_n)\ , \tag{8.7}$$

$\forall x_1, x_2, \ldots, x_n$. *Find all such functions* f, f_i*, none of which vanishes identically.*

Comment. Equation (8.7) is known as the **second Pexider functional equation** or the **exponential Pexider functional equation**.

Solution. Setting $x_1 = x_2 = \cdots = x_n = 0$ in the defining equation, we see that

$$f(0) \;=\; f_1(0)\, f_2(0) \ldots f_n(0)\ . \tag{8.8}$$

Also, we set successively $x_1 = x_2 = \cdots = x_{i-1} = x_{i+1} = \cdots = x_n = 0$, for $i = 1, 2, \ldots, n$:

$$\begin{aligned}
f(x_1) \;&=\; f_1(x_1)\, f_2(0)\, f_3(0) \ldots f_{n-1}(0)\, f_n(0)\ ,\\
f(x_2) \;&=\; f_1(0)\, f_2(x_2)\, f_3(0) \ldots f_{n-1}(0)\, f_n(0)\ ,\\
\cdots \;&\quad\; \cdots\\
f(x_n) \;&=\; f_1(0)\, f_2(0)\, f_3(0) \ldots f_{n-1}(0)\, f_n(x_n)\ ,
\end{aligned}$$

Multiplying the last n equations and using the defining equation and (8.8), we find

$$f(x_1)\, f(x_2) \ldots f(x_n) \;=\; f(x_1 + x_2 + \cdots + x_n)\, f(0)^{n-1}\ .$$

If $f(0) = 0$, then $f_1(0)\, f_2(0) \dots f_n(0) = 0$, so at least one of $f_1(0), f_2(0), \dots, f_n(0)$ will vanish, which in turn implies that $f(x) = 0$, $\forall x \in \mathbb{R}$. Since we are looking for solutions which include no identically vanishing functions, $f(0) \neq 0$ and

$$\frac{f(x_1)}{f(0)} \cdots \frac{f(x_n)}{f(0)} = \frac{f(x_1 + x_2 + \cdots + x_n)}{f(0)} .$$

That is, the function $f(x)/f(0)$ satisfies the equation (5.10), equivalent to the exponential Cauchy equation (5.8), and therefore f is given by

$$f(x) = \left(\prod_{i=1}^{n} a_i \right) c^x , \quad c \in \mathbb{R}_+^* ,$$

where we set $f_i(0) = a_i$, $\forall i$. Then

$$f_i(x) = a_i\, c^x , \quad \forall i = 1, 2, \dots, n . \qquad \square$$

8.3 Third Pexider Equation

Problem 8.4. *Let $f, f_i : \mathbb{R}_+^* \to \mathbb{R}$, $i = 1, 2, \dots, n$, be $n + 1$ continuous functions that satisfy the functional relation*

$$f(x_1\, x_2 \dots x_n) = f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n) , \tag{8.9}$$

$\forall x_1, x_2, \dots, x_n$. Find all such non-constant functions f, f_i.

Comment. Equation (8.9) is known as the **third Pexider functional equation** or the **logarithmic Pexider functional equation** .

Solution. Setting $x_1 = x_2 = \cdots = x_n = 1$ in the defining equation, we see that

$$f(1) = f_1(1) + f_2(1) + \cdots + f_n(1) . \tag{8.10}$$

Also, we set successively $x_1 = x_2 = \cdots = x_{i-1} = x_{i+1} = \cdots = x_n = 1$, for $i = 1, 2, \dots, n$:

$$\begin{aligned}
f(x_1) &= f_1(x_1) + f_2(1) + f_3(1) + \cdots + f_{n-1}(1) + f_n(1) , \\
f(x_2) &= f_1(1) + f_2(x_2) + f_3(1) + \cdots + f_{n-1}(1) + f_n(1) , \\
\cdots & \quad \cdots \\
f(x_n) &= f_1(1) + f_2(1) + f_3(1) + \cdots + f_{n-1}(1) + f_n(x_n) ,
\end{aligned}$$

Adding the last n equations and using the defining equation and (8.10), we find

$$f(x_1) + f(x_2) + \cdots + f(x_n) = f(x_1\, x_2 \dots x_n) + (n - 1)\, f(1) ,$$

or

$$[f(x_1) - f(1)] + \cdots + [f(x_n) - f(1)] = [f(x_1\, x_2 \dots x_n) - f(1)] .$$

That is, the function $f(x) - f(1)$ satisfies equation (5.16), equivalent to the logarithmic Cauchy equation (5.14), and it is therefore either identically zero or $\log_\gamma x$, for some positive γ. Therefore the non-constant f is given by

$$f(x) = \log_\gamma x + \sum_{i=1}^{n} a_i ,$$

where we set $f_i(1) = a_i$, $\forall i$. Then

$$f_i(x) = \log_\gamma x + a_i , \quad \forall i = 1, 2, \ldots, n . \qquad \square$$

8.4 Fourth Pexider Equation

Problem 8.5. *Let $f, f_i : \mathbb{R}_+^* \to \mathbb{R}$, $i = 1, 2, \ldots, n$, be $n + 1$ continuous functions that satisfy the functional relation*

$$f(x_1 x_2 \ldots x_n) = f_1(x_1) f_2(x_2) \ldots f_n(x_n) , \tag{8.11}$$

$\forall x_1, x_2, \ldots, x_n$. *Find all such non-constant functions f, f_i.*

Comment. Equation (8.11) is known as the **fourth Pexider functional equation** or the **power Pexider functional equation**.

Solution. Setting $x_1 = x_2 = \cdots = x_n = 1$ in the defining equation, we see that

$$f(1) = f_1(1) f_2(1) \ldots f_n(1) . \tag{8.12}$$

Also, we set successively $x_1 = x_2 = \cdots = x_{i-1} = x_{i+1} = \cdots = x_n = 1$, for $i = 1, 2, \ldots, n$:

$$\begin{aligned} f(x_1) &= f_1(x_1) f_2(1) f_3(1) \ldots f_{n-1}(1) f_n(1) , \\ f(x_2) &= f_1(1) f_2(x_2) f_3(1) \ldots f_{n-1}(1) f_n(1) , \\ \ldots & \quad \ldots \\ f(x_n) &= f_1(1) f_2(1) f_3(1) \ldots f_{n-1}(1) f_n(x_n) . \end{aligned}$$

Multiplying the last n equations and using the defining equation and (8.12), we find

$$f(x_1) f(x_2) \ldots f(x_n) = f(x_1 x_2 \ldots x_n) f(1)^{n-1} .$$

If $f(1) = 0$, then $f_1(1) f_2(1) \ldots f_n(1) = 0$. At least one of $f_1(1), f_2(1), \ldots, f_n(1)$ will vanish, which in turn implies that $f(x) = 0$, $\forall x \in \mathbb{R}_+^*$. Since we are looking for solutions which include no identically vanishing functions, $f(1) \neq 0$ and

$$\frac{f(x_1)}{f(1)} \ldots \frac{f(x_n)}{f(1)} = \frac{f(x_1 x_2 \ldots x_n)}{f(1)} .$$

That is, the function $f(x)/f(1)$ satisfies equation (5.22), equivalent to the power Cauchy equation, (5.20), and it is therefore either identically zero or x^c, for a constant c. Therefore the non-constant f is given by

$$f(x) = \left(\prod_{i=1}^{n} a_i\right) x^c ,$$

where we set $f_i(1) = a_i$, $\forall i$. Then

$$f_i(x) = a_i\, x^c , \quad \forall i = 1, 2, \ldots, n . \qquad \square$$

8.5 Vincze Functional Equations

A further way to generalize the equations

$$\begin{aligned} f(x+y) &= g(x) + h(y) , \\ f(x+y) &= g(x)\, h(y) , \end{aligned}$$

is the functional equation

$$f(x+y) = g_1(x)\, g_2(y) + h(y) .$$

Similarly we can generalize the equations

$$\begin{aligned} f(x\,y) &= g(x) + h(y) , \\ f(x\,y) &= g(x)\, h(y) , \end{aligned}$$

to the functional equation

$$f(x\,y) = g_1(x)\, g_2(y) + h(y) .$$

In this section we solve such equations.

Problem 8.6. *Let $f, g_1, g_2, h : \mathbb{R} \to \mathbb{R}$, be continuous functions that satisfy the functional relation*

$$f(x+y) = g_1(x)\, g_2(y) + h(y) , \tag{8.13}$$

$\forall x, y$. Find all such functions.

Comment. Equation (8.13) is known as the **first Vincze** (Vincze I) **functional equation**.

Solution. Setting $y = 0$ in the defining equation (8.13), we see that

$$f(x) = g_1(x)\, g_2(0) + h(0) . \tag{8.14}$$

At this point we examine two separate cases: $g_2(0) = 0$ or $g_2(0) \neq 0$.

Case I: If $g_2(0) = 0$, then $f(x) = h(0)$, for all x. For $x = 0$, equation (8.13) then gives $h(y) = h(0) - g_1(0)g_2(y)$. Substituting this expression back in (8.13), we find $g_2(y)\,[g_1(x) - g_1(0)] = 0$. Therefore, either $g_2(x) = 0$ for all x or $g_1(x) = g_1(0)$ for all x. Therefore, the solutions of (8.13) in this case are:

$$\left\{\begin{array}{lcl} f(x) & = & a \\ g_1(x) & = & \text{arbitrary} \\ g_2(x) & = & 0 \\ h(x) = a & & \end{array}\right\} \quad \text{or} \quad \left\{\begin{array}{lcl} f(x) & = & a \\ g_1(x) & = & b \\ g_2(x) & = & \text{arbitrary}\,,\, g_2(0) = 0 \\ h(x) = a - b g_2(x) & & \end{array}\right\}.$$

Case II: If $g_2(0) \neq 0$, then equation (8.14) can be solved for $g_1(x)$:

$$g_1(x) \;=\; \frac{f(x) - h(0)}{g_2(0)}\,. \tag{8.15}$$

Substituting this result in (8.13), we find

$$f(x+y) \;=\; f(x)\,\frac{g_2(y)}{g_2(0)} + h(y) - h(0)\,\frac{g_2(y)}{g_2(0)}\,.$$

We can rewrite this equation in a nicer form using the functions

$$\tilde{g}(x) \;=\; \frac{g_2(x)}{g_2(0)}\,, \qquad H(y) \;=\; h(y) - h(0)\,\frac{g_2(y)}{g_2(0)}\,.$$

Then

$$f(x+y) \;=\; f(x)\,\tilde{g}(y) + H(y)\,. \tag{8.16}$$

We have succeeded to reduce the initial equation (8.13) with four unknown functions to an equation (8.16) with three unknown functions. We can continue this reduction until we find an equation with only one unknown function. Setting $x = 0$ in (8.16) we find $H(y) = f(y) - f(0)\tilde{g}(y)$. When this is inserted back into (8.16)

$$f(x+y) \;=\; [f(x) - f(0)]\,\tilde{g}(y) + f(y)\,,$$

i.e. we discover an equation with two unknown functions. It is more convenient to work with the function

$$F(x) \;=\; f(x) - f(0)\,.$$

Therefore

$$F(x+y) \;=\; F(x)\,\tilde{g}(y) + F(y)\,. \tag{8.17}$$

We can study now two separate subcases: either $\tilde{g}(y) \equiv 1$ or $\tilde{g}(y) \not\equiv 1$.

Subcase IIa: $\tilde{g}(y) \equiv 1$ or equivalently $g_2(x) \equiv g_2(0)$. In this case, $F(x)$ satisfies (5.1):

$$F(x+y) \;=\; F(x) \,+ F(y)\,,$$

with solution $F(x) = c\,x$ or $f(x) = c\,x + f(0)$. Equation (8.15) then gives $g_1(x) = \frac{f(0)-h(0)}{g_2(0)} + \frac{c}{g_2(0)}\,x$ and (8.13) that $h(x) = c\,x + h(0)$. Let's list the complete solution:

$$\left\{ \begin{array}{l} f(x) \;=\; c\,x + a \\ g_1(x) \;=\; \frac{c}{b}\,x + \frac{a-d}{b} \\ g_2(x) \;=\; b \\ h(x) = c\,x + d \end{array} \right\} .$$

<u>Subcase IIb</u>: $\tilde{g}(y) \not\equiv 1$. Then there exists $y_0 \in \mathbb{R}^*$ such that $\tilde{g}(y) \neq 1$. In equation (8.17), set $y = y_0$:

$$F(x + y_0) \;=\; F(x)\,\tilde{g}(y_0) + F(y_0)\,.$$

In the same equation we now set $x = y_0$ and $y = x$:

$$F(y_0 + x) \;=\; F(y_0)\,\tilde{g}(x) + F(x)\,.$$

Subtracting the last two equations we find

$$F(x) \;=\; c\,[\tilde{g}(x) - 1]\,, \tag{8.18}$$

where $c = F(y_0)/[\tilde{g}(y_0) - 1]$. We now have two new subcases $c = 0$ and $c \neq 0$.

<u>Subsubcase IIb1</u>: $c = 0$ and therefore $f(x) = f(0)$ and $g_1(x) = [f(0) - h(0)]/g_2(0)$. The function $g_2(x)$ is arbitrary, except for $g_2(0) \neq 0$ and that there is at least one point y_0 where it has value other from $g_2(0)$ and $h(x) = f(0) - g_2(x)\,[f(0) - h(0)]/g_2(0)$.

$$\left\{ \begin{array}{l} f(x) \;=\; a \\ g_1(x) \;=\; b \\ g_2(x) \;=\; \text{arbitrary}\,, \quad g_2(0) \neq 0\,, \quad g_2(y_0) \neq g_2(0) \\ h(x) = a - b\,g_2(x) \end{array} \right\} .$$

<u>Subsubcase IIb2</u>: $c \neq 0$. Substituting (8.18) in (8.17) we find

$$\tilde{g}(x + y) \;=\; \tilde{g}(x)\,\tilde{g}(y)\,,$$

i.e. the function $\tilde{g}(x)$ satisfies the exponential Cauchy functional equation (5.8). So either $\tilde{g}(x) \equiv 0$ (which is rejected since it does not satisfy the condition $\tilde{g}(0) = 1$) or $\tilde{g}(x) = a^x$ for some positive constant a. Therefore $g_2(x) = g_2(0)\,a^x$ and $f(x) = c\,a^x - c + f(0)$. From (8.15), $g_1(x) = [c/g_2(0)]\,a^x + (f(0) - h(0) - c)/g_2(0)$. Finally, $h(x) = (f(0) - c) - (f(0) - c - h(0))\,a^x$. Let's record the final result:

$$\left\{ \begin{array}{l} f(x) \;=\; c\,a^x + k \\ g_1(x) \;=\; \frac{c}{d}\,a^x + \frac{b}{d} \\ g_2(x) \;=\; d\,a^x \\ h(x) = k - b\,a^x \end{array} \right\} . \qquad \square$$

Problem 8.7. *: Let $f, g_1, g_2, h : \mathbb{R} \to \mathbb{R}$ be continuous functions that satisfy the functional relation*

$$f(x\,y) \;=\; g_1(x)\,g_2(y) + h(y)\,, \tag{8.19}$$

$\forall x, y$. Find all such functions.

Comment. Equation (8.19) is known as the **second Vincze** (Vincze II) **functional equation**.

Solution. Setting $y = 1$ in the defining equation (8.19), we see that

$$f(x) \;=\; g_1(x)\,g_2(1) + h(1)\,. \tag{8.20}$$

At this point we examine two separate cases: $g_2(1) = 0$ or $g_2(1) \neq 0$.

Case I: If $g_2(1) = 0$, then $f(x) = h(1)$, for all x. For $x = 1$, equation (8.19) then gives $h(y) = h(1) - g_1(1)g_2(y)$. Substituting this expression back in (8.19), we find $g_2(y)\,[g_1(x) - g_1(1)] = 0$. Therefore, either $g_2(x) = 0$ for all x or $g_1(x) = g_1(1)$ for all x. Therefore, the solutions of (8.19) in this case are:

$$\left\{\begin{array}{l} f(x) \;=\; a \\ g_1(x) \;=\; \text{arbitrary} \\ g_2(x) \;=\; 0 \\ h(x) = a \end{array}\right\} \quad \text{or} \quad \left\{\begin{array}{l} f(x) \;=\; a \\ g_1(x) \;=\; b \\ g_2(x) \;=\; \text{arbitrary}\,, g_2(1) = 0 \\ h(x) = a - b g_2(x) \end{array}\right\}.$$

Case II: If $g_2(1) \neq 0$, then equation (8.20) can be solved for $g_1(x)$:

$$g_1(x) \;=\; \frac{f(x) - h(1)}{g_2(1)}\,. \tag{8.21}$$

Substituting this result in (8.19), we find

$$f(x\,y) \;=\; f(x)\,\frac{g_2(y)}{g_2(1)} + h(y) - h(1)\,\frac{g_2(y)}{g_2(1)}\,.$$

We can rewrite in a nicer form using the functions

$$\tilde{g}(x) \;=\; \frac{g_2(x)}{g_2(1)}\,, \qquad H(x) \;=\; h(y) - h(1)\,\frac{g_2(y)}{g_2(1)}\,.$$

Then

$$f(x\,y) \;=\; f(x)\,\tilde{g}(y) + H(y)\,. \tag{8.22}$$

We have succeeded to reduce the initial equation (8.19) with four unknown functions to an equation (8.22) with three unknown functions. We can continue this reduction until we find an equation with only one unknown function. Setting $x = 1$ in (8.22) we find $H(y) = f(y) - f(1)\tilde{g}(y)$. When this is inserted back in (8.22)

$$f(x\,y) \;=\; [f(x) - f(1)]\,\tilde{g}(y) + f(y)\,,$$

i.e. we discover an equation with two unknown functions. It is more convenient to work with the function

$$F(x) \;=\; f(x) - f(1)\,.$$

Therefore

$$F(x\,y) \;=\; F(x)\,\tilde{g}(y) + F(y)\,. \tag{8.23}$$

We can study now two separate subcases: either $\tilde{g}(y) \equiv 1$ or $\tilde{g}(y) \not\equiv 1$.

Subcase IIa: $\tilde{g}(y) \equiv 1$ or equivalently $g_2(x) \equiv g_2(1)$. In this case, $F(x)$ satisfies (5.14):

$$F(x\,y) \;=\; F(x) \,+ F(y)\,,$$

with solution $F(x) = \log_\gamma x$ or $f(x) = \log_\gamma x + f(1)$. Equation (8.21) then gives $g_1(x) = \frac{f(1)-h(1)}{g_2(1)} + \frac{1}{g_2(1)} \log_\gamma x$ and (8.19) that $h(x) = \log_\gamma x + h(1)$. Let's list the complete solution:

$$\left\{\begin{array}{l} f(x) \;=\; \log_\gamma x + a \\ g_1(x) \;=\; \frac{1}{b}\log_\gamma x + \frac{a-d}{b} \\ g_2(x) \;=\; b \\ h(x) = \log_\gamma x + d \end{array}\right\}\,.$$

Subcase IIb: $\tilde{g}(y) \not\equiv 1$. Then there exists $y_0 \in \mathbb{R}^*$ such that $\tilde{g}(y) \neq 1$. In equation (8.23), set $y = y_0$:

$$F(x\,y_0) \;=\; F(x)\,\tilde{g}(y_0) + F(y_0)\,.$$

In the same equation we now set $x = y_0$ and $y = x$:

$$F(y_0\,x) \;=\; F(y_0)\,\tilde{g}(x) + F(x)\,.$$

Subtracting the last two equations we find

$$F(x) \;=\; c\,[\tilde{g}(x) - 1]\,, \tag{8.24}$$

where $c = F(y_0)/[\tilde{g}(y_0) - 1]$. We now have two new subcases $c = 0$ and $c \neq 0$.

Subsubcase IIb1: $c = 0$ and therefore $f(x) = f(1)$ and $g_1(x) = [f(1) - h(1)]/g_2(1)$. The function $g_2(x)$ is arbitrary, except for $g_2(0) \neq 0$ and that there is at least one point y_0 where it has value other from $g_2(1)$, and $h(x) = f(1) - g_2(x)\,[f(1) - h(1)]/g_2(1)$.

$$\left\{\begin{array}{l} f(x) \;=\; a \\ g_1(x) \;=\; b \\ g_2(x) \;=\; \text{arbitrary}\,, \quad g_2(1) \neq 0\,, \quad g(y_0) \neq g_2(1) \\ h(x) = a - b\,g_2(x) \end{array}\right\}\,.$$

Subsubcase IIb2: $c \neq 0$. Substituting (8.24) in (8.23) we find

$$\tilde{g}(x\,y) \;=\; \tilde{g}(x)\,\tilde{g}(y)\,,$$

i.e. the function $\tilde{g}(x)$ satisfies the power Cauchy functional equation (5.20). So either $\tilde{g}(x) \equiv 0$ (which is rejected since does not satisfy the condition $\tilde{g}(1) = 1$) or $\tilde{g}(x) = x^a$ for some constant a. Therefore $g_2(x) = g_2(1)\, x^a$ and $f(x) = c\, x^a - c + f(1)$. From (8.21), $g_1(x) = [c/g_2(1)]\, a^x + (f(1) - h(1) - c)/g_2(1)$. Finally, $h(x) = (f(1) - c) - (f(1) - c - h(1))\, a^x$. Let's record the final result:

$$\left\{ \begin{array}{l} f(x) \ = \ c\, x^a + k \\ g_1(x) \ = \ \frac{c}{d}\, x^a + \frac{b}{d} \\ g_2(x) \ = \ d\, x^a \\ h(x) = k - b\, x^a \end{array} \right\} . \qquad \square$$

8.6 Wilson Functional Equations

Wilson has generalized the (7.21) equation in a similar fashion with Pexider's generalization of the Cauchy equations.

Problem 8.8. *: Let $f, g : \mathbb{R} \to \mathbb{R}$ be continuous functions that satisfy the functional equation*

$$f(x + y) + f(x - y) \ = \ 2\, f(x)\, g(y)\ , \quad \forall x, y \in \mathbb{R}\ . \tag{8.25}$$

Find all functions f, g.

Comment. We will call equation (8.25) the **first Wilson** (Wilson I) **functional equation**. □

We will not present the solution as it is relative long. Instead we shall wait until Chapter 12 where stronger than continuity conditions may be imposed on the function which simplify the problem considerably. Here, we are only presenting the answer.

Answer (Wilson). The most general continuous solutions of (8.25) are:

1. $f(x) \ \equiv \ 0$, $g(x) \ = \ \text{arbitrary}$;
2. $f(x) \ = \ a\cos(\lambda x) + b\sin(\lambda x)$, $g(x) \ = \ \cos(\lambda x)$;
3. $f(x) \ = \ a\cosh(\lambda x) + b\sinh(\lambda x)$, $g(x) \ = \ \cos(\lambda x)$;
4. $f(x) \ = \ a + b\, x$, $g(x) \ = \ 1$. □

Using this result, Wilson generalized the equation further:

Problem 8.9. *Let $e, f, g, h : \mathbb{R} \to \mathbb{R}$ be continuous functions that satisfy the functional equation*

$$e(x + y) + f(x - y) \ = \ 2\, h(x)\, g(y)\ , \quad \forall x, y \in \mathbb{R}\ . \tag{8.26}$$

Find all functions e, f, g, k.

Comment. We will call equation (8.26) the **second Wilson** (Wilson II) **functional equation**. □

Again, we only present the solution to this equation.

Answer (Wilson). The most general continuous solutions of (8.26) are:

1. $e(x) = c$, $f(x) = -c$, $h(x) = 0$, $g(x) =$ arbitrary;
2. $e(x) = c$, $f(x) = -c$, $h(x) =$ arbitrary, $g(x) = 0$;
3. $e(x) = a\cos(\lambda x) + b\sin(\lambda x) + s$, $f(x) = c\cos(\lambda x) + d\sin(\lambda x) - s$, $g(x) = A\cos(\lambda x) + B\sin(\lambda x)$, $h(x) = C\cos(\lambda x) + D\sin(\lambda x)$ with $a + c = 2AC$, $c - a = 2BD$, $b + d = 2AD$, $b - d = 2BC$;
4. $e(x) = a\cosh(\lambda x) + b\sinh(\lambda x) + s$, $f(x) = c\cosh(\lambda x) + d\sinh(\lambda x) - s$, $g(x) = A\cosh(\lambda x) + B\sinh(\lambda x)$, $h(x) = C\cosh(\lambda x) + D\sinh(\lambda x)$ with $a + c = 2AC$, $a - c = 2BD$, $b + d = 2AD$, $b - d = 2BC$;
5. $e(x) = ax^2 + bx + c$, $f(x) = -ax^2 + b'x + c'$, $g(x) = Ax + B$, $h(x) = Cx + D$ with $a = AC/2$, $b + b' = 2BC$, $b - b' = 2AD$, $c + c' = 2BD$. □

8.7 Solved Problems

As in the case of the Cauchy functional equations, many functional equations can be solved by reducing them to one of the Pexider functional equations. In this section, we present several examples.

Problem 8.10. *Find the continuous solutions* $f, g, h : \mathbb{R} \to \mathbb{R}$ *of the functional equation*

$$f\left(\frac{x+y}{2}\right) = \frac{g(x) + h(y)}{2}, \quad \forall x, y \in \mathbb{R}.$$

Solution. We define the continuous function

$$F(x) = 2f\left(\frac{x}{2}\right).$$

Then the functional equation takes the form

$$F(x+y) = g(x) + h(y),$$

which has the solution

$$F(x) = cx + a + b, \quad g(x) = cx + a, \quad h(x) = cx + b,$$

where a, b, c are constants. Therefore

$$f(x) = cx + \frac{a+b}{2}, \quad g(x) = cx + a, \quad h(x) = cx + b. \qquad \square$$

Problem 8.11. *Find the continuous solutions* $f, g, h : \mathbb{R} \to \mathbb{R}$ *of the functional equation*

$$f(x+y) = \frac{g(x)+h(y)}{1-\frac{g(x)\,h(y)}{c^2}}, \quad \forall x, y \in \mathbb{R}, \tag{8.27}$$

where c is a constant.

Solution. We define the new functions $F(x), G(x), H(x)$ by

$$f(x) = c\,\tan(F(x)), \quad g(x) = c\,\tan(G(x)), \quad h(x) = c\,\tan(H(x)),$$

in terms of which the given functional equation takes the form

$$\tan(F(x+y)) = \tan(G(x)+H(y)),$$

or

$$F(x+y) = G(x)+H(y)+m\,\pi, \quad m \in \mathbb{Z}.$$

Rearranging the terms of the equation, we find

$$[F(x+y)-m\pi] = [G(x)-m\pi]+[H(y)-m\,\pi],$$

which is of course equation (8.5) with solutions

$$F(x) = k\,x+a+b+m\pi, \quad G(x) = k\,x+a+m\pi, \quad H(x) = k\,x+b+m\pi.$$

Therefore

$$f(x) = c\,\tan(k\,x+a+b), \quad g(x) = c\,\tan(k\,x+a), \quad h(x) = c\,\tan(k\,x+b).$$ □

Comment. It should now be evident to the reader how to solve the functional equations

$$f(x+y) = \frac{f(x)+f(y)}{1+\frac{f(x)\,f(y)}{c^2}}, \tag{8.28}$$

$$f(x+y) = \frac{f(x)\,f(y)-1}{f(x)\,f(y)},$$

(which admit the solutions $f(x) = c\tanh kx$, $f(x) = \cot kx$ respectively) and their generalizations to three functions. □

Comment. This problem, for a single function only (and $c = 1$),

$$f(x+y) = \frac{f(x)+f(y)}{1-\frac{f(x)\,f(y)}{c^2}}, \tag{8.29}$$

had been proposed by Hungary for the IMO 1977. (See IMO 1977 longlist in [16].) However, the selection committee did not consider it as a viable candidate for the competition. (See corresponding shortlist in [16].) Most probably, the committee's decision was based on the fact that the problem was not very original! Instead, it was widely known — at least to physicists. The functional equation (8.28) is nothing else but the relativistic addition of velocities:

$$v_{rel} = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}},$$

where c is the speed of light. Physicists routinely use (and had used well before 1977) the transformation

$$v = c \tanh\theta,$$

since the law becomes additive in the variable θ (which is known as the **rapidity**). Using analytic continuation $\theta \mapsto i\theta$, equation (8.28) is mapped to equation (8.29). Perhaps, equations of the form (8.27) and (8.28) should be called the *Einstein-Lorentz equations* to honor the founders of Special Relativity. □

Problem 8.12. *Find the continuous solutions $f, g, h : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(\sqrt[n]{x^n + y^n}) = g(x)\,h(y), \quad \forall x, y \in \mathbb{R},$$

where n is a constant.

Solution. For $y = 0$, the defining equation gives $f(x) = g(x)h(0)$. If $h(0) = 0$, then $f(x) \equiv 0$ and $g(x)$ is arbitrary.

For $x = 0$, the defining equation gives $f(y) = g(0)h(y)$. If $g(0) = 0$, then $f(x) \equiv 0$ and $h(x)$ is arbitrary.

We shall look for solutions in which none of the functions f, g, h vanishes identically. In this case, we define the new functions F, G, H by

$$\begin{aligned} F(u) &= f(\sqrt[n]{u}), \\ G(u) &= g(\sqrt[n]{u}), \\ H(u) &= h(\sqrt[n]{u}). \end{aligned}$$

Then the defining equation becomes

$$F(u + v) = G(u)\,H(v),$$

where $u = x^n$, $v = y^n$. The solution of this equation with all function F, G, H non-vanishing is

$$F(u) = ab\,c^u, \quad G(u) = a\,c^u, \quad H(u) = b\,c^u,$$

with a, b any constants and c a positive constant. Then

$$f(x) = ab\,c^{x^n}, \quad g(x) = a\,c^{x^n}, \quad h(x) = b\,c^{x^n}. \qquad \square$$

Chapter 9

Vector and Matrix Variables

So far we have studied functional equations that include functions $f(x)$ of one independent variable x, i.e. the domain of f is a subset of $\mathbb{R}$. However, one can study equations that include functions $f(x_1, x_2, \ldots, x_n)$ of many independent variables $x_1, x_2, \ldots, x_n$, i.e. the domain of f is a subset of $\mathbb{R}^n$. One can view the collection of the independent variables as a vector $\vec{x} = (x_1, x_2, \ldots, x_n)$. Examples of equations of this category are

$$\begin{aligned} f(\vec{x}+\vec{y}) &= f(\vec{x}) + f(\vec{y})\,, \\ f(\vec{x}+\vec{y}) &= f(\vec{x})\, f(\vec{y})\,, \\ f(\vec{x},\vec{z}) &= f(\vec{x},\vec{y}) + f(\vec{y},\vec{z})\,. \end{aligned}$$

Even more, it is possible that we study functions that take values in $\mathbb{R}^m$. Obviously, in these cases the function itself is a vector. Examples of equations of this category are

$$\begin{aligned} \vec{f}(x\,y) &= \vec{f}(x) + \vec{f}(y)\,, \\ f(x\,y) &= \vec{g}(x) \cdot \vec{g}(y)\,, \\ \vec{f}(\vec{x}+\vec{y}) &= \vec{g}(\vec{x}) + \vec{h}(\vec{y})\,, \\ f(\vec{x}+\vec{y}) &= \vec{g}(\vec{x}) \cdot \vec{h}(\vec{y})\,, \\ \vec{f}(\vec{x},\vec{z}) &= \vec{f}(\vec{x},\vec{y}) + \vec{f}(\vec{y},\vec{z})\,. \end{aligned}$$

Often it is useful to consider functions from a subset of all matrices to another subset of all matrices. For example, let $\mathcal{M}(n)$ be the set of all non-singular square matrices of order n with real elements and $f : \mathcal{M}(n) \to \mathbb{R}$ such that

$$f(\boldsymbol{X}\,\boldsymbol{Y}) = f(\boldsymbol{X})\, f(\boldsymbol{Y})\,,$$

for all $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{M}(n)$. Another example is the functional equation

$$F(\boldsymbol{X}\,\boldsymbol{Y}) = F(\boldsymbol{X})\, F(\boldsymbol{Y})\,,$$

for the function $F : \mathcal{M}(n) \to \mathcal{M}(m)$.

9.1 Cauchy & Pexider Type Equations

Problem 9.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function that satisfies the functional relation*

$$f(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \;=\; f(x_1, x_2, \ldots, x_n) + f(y_1, y_2, \ldots, y_n)\,.$$

Find all functions f that satisfy the above condition.

Solution. In the defining equation we set $x_1 = \cdots = \hat{x}_k = \cdots = x_n = 0$ and $y_1 = \cdots = \hat{y}_k = \cdots = y_n = 0$ (where the hat above a variable means omission of the corresponding variable). Then

$$f(0, \ldots, 0, x_k + y_k, 0 \ldots, 0) \;=\; f(0, \ldots, 0, x_k, 0 \ldots, 0) + f(0, \ldots, 0, y_k, 0 \ldots, 0)\,.$$

Let $f_k(x) = f(0, \ldots, 0, x, 0 \ldots, 0)$. Then the previous equation reads

$$f_k(x + y) \;=\; f_k(x) + f_k(y)\,,$$

with continuous solution $f_k(x) = c_k\,x$, and c_k a constant.

We can express the function $f(x_1, x_2, \ldots, x_n)$ in terms of $f_k(x)$ as follows:

$$\begin{aligned}
f(x_1, x_2, \ldots, x_n) &= f(x_1 + 0, 0 + x_2, \ldots, 0 + x_n) \\
&= f(x_1, 0, \ldots, 0) + f(0, x_2, \ldots, x_n) \\
&= f(x_1, 0, \ldots, 0) + f(0 + 0, x_2 + 0, 0 + x_3, \ldots, 0 + x_n) \\
&= f(x_1, 0, \ldots, 0) + f(0, x_2, 0, \ldots, 0) + f(0, 0, x_3, \ldots, x_n) \\
&= \ldots\ldots \\
&= f(x_1, 0, \ldots, 0) + f(0, x_2, 0, \ldots, 0) + \cdots + f(0, 0, \ldots, 0, x_n) \\
&= f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n)\,.
\end{aligned}$$

Therefore

$$f(x_1, x_2, \ldots, x_n) \;=\; c_1\,x_1 + c_2\,x_2 + \cdots + c_n\,x_n\,. \qquad \square$$

Comment. Above we used 'expanded' notation. In more compact notation, we just proved that the only solution of the functional equation

$$f(\vec{x} + \vec{y}) \;=\; f(\vec{x}) + f(\vec{y})\,.$$

is the linear function

$$f(\vec{x}) \;=\; \vec{c} \cdot \vec{x}\,,$$

where $\vec{c}$ is a constant vector. $\square$

The following problem is a direct consequence of the previous one.

Problem 9.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous symmetric function*[1] *that satisfies the functional equation*

$$f(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) = f(x_1, x_2, \ldots, x_n) + f(y_1, y_2, \ldots, y_n) ,$$

and is such that

$$f(x, x, \ldots, x) = x .$$

Then f is the arithmetic mean of the variables

$$f(x_1, x_2, \ldots, x_n) = \frac{x_1 + x_2 + \cdots + x_n}{n} .$$

Solution. Since $f(x_1, x_2, \ldots, x_n)$ is a symmetric function, all functions $f_k(x)$ defined previously are identical; let's use the common symbol $g(x)$ to indicate them. Then

$$f(x_1, x_2, \ldots, x_n) = g(x_1) + g(x_2) + \cdots + g(x_n) ,$$

and

$$x = f(x, x, \ldots, x) = n\, g(x) \Rightarrow g(x) = \frac{x}{n} .$$

From this the result follows immediately. □

Problem 9.3. *Let $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$ be a continuous function that satisfies the functional relation*

$$\vec{f}(\vec{x} + \vec{y}) = \vec{f}(\vec{x}) + \vec{f}(\vec{y}) .$$

Find all functions $\vec{f}$ that satisfy the above conditions.

Solution. Let $f_\ell(\vec{x})$, $\ell = 1, 2, \ldots, m$ be the components of the function $\vec{f}$. The given vector equation results in m scalar equations

$$f_\ell(\vec{x} + \vec{y}) = f_\ell(\vec{x}) + f_\ell(\vec{y}) .$$

The continuous solutions of these equations were found above:

$$f_\ell(\vec{x}) = \vec{c}_\ell \cdot \vec{x} ,$$

where $\vec{c}_\ell$ is a constant vector.

It might be more appealing to write the above equation in another form:

$$\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} \vec{c}_1 \\ \vec{c}_2 \\ \vdots \\ \vec{c}_n \end{bmatrix} \cdot \vec{x} .$$

In other words, we have arranged the constant vectors in a matrix $\boldsymbol{C}$ such that $\vec{f} = \boldsymbol{C} \cdot \vec{x}$. □

[1] A symmetric function is a function that remains invariant under any permutation of its variables. (See page 188 for the corresponding definition of symmetric polynomials.)

Problem 9.4. *Let* $\vec{f}, \vec{g}, \vec{h} : \mathbb{R}^n \to \mathbb{R}^m$ *be continuous functions that satisfy the functional equation*

$$\vec{f}(\vec{x} + \vec{y}) = \vec{g}(\vec{x}) + \vec{h}(\vec{y}) \tag{9.1}$$

Find all functions $\vec{f}, \vec{g}, \vec{h}$ *that satisfy the above conditions.*

Solution. Setting $\vec{x} = \vec{0}$ and $\vec{g}(\vec{0}) = \vec{a}$ we get:

$$\vec{f}(\vec{y}) = \vec{h}(\vec{y}) + \vec{a}\,.$$

Repeating the process for $\vec{y}$ and setting $\vec{h}(\vec{0}) = \vec{b}$ we get:

$$\vec{f}(\vec{x}) = \vec{g}(\vec{x}) + \vec{b}\,.$$

Rearranging terms we have:

$$\begin{aligned} \vec{h}(\vec{y}) &= \vec{f}(\vec{y}) - \vec{a}\,, \\ \vec{g}(\vec{y}) &= \vec{f}(\vec{y}) - \vec{b}\,. \end{aligned}$$

Plugging this into the defining relation gives us:

$$\vec{f}(\vec{x} + \vec{y}) = \vec{f}(\vec{x}) + \vec{f}(\vec{y}) - \vec{a} - \vec{b}\,.$$

Next we define $\vec{j}(\vec{x}) = \vec{f}(\vec{x}) - \vec{a} - \vec{b}$, which gives us the relation:

$$\vec{j}(\vec{x} + \vec{y}) = \vec{j}(\vec{x}) + \vec{j}(\vec{y})\,,$$

the solution of which is

$$\vec{j}(\vec{x}) = \boldsymbol{C} \cdot \vec{x}\,,$$

where $\boldsymbol{C}$ is a constant matrix. Thus the general solution to (9.1) is:

$$\begin{aligned} \vec{f}(\vec{x}) &= \boldsymbol{C} \cdot \vec{x} + \vec{a} + \vec{b}\,, \\ \vec{h}(\vec{x}) &= \boldsymbol{C} \cdot \vec{x} + \vec{b}\,, \\ \vec{g}(\vec{x}) &= \boldsymbol{C} \cdot \vec{x} + \vec{a}\,. \end{aligned}$$

□

9.2 Solved Problems

Problem 9.5 (Putnam 1959)**.** *For each positive integer* n*, let* f_n *be a real-valued symmetric function of* n *variables. Suppose that for all* n *and for all real numbers* $x_1, x_2, \ldots, x_{n+1}, y$ *it is true that*
(a) $f_n(x_1 + y, x_2 + y, \ldots, x_n + y) = f_n(x_1, x_2, \ldots, x_n) + y$;
(b) $f_n(-x_1, -x_2, \ldots, -x_n) = -f_n(x_1, x_2, \ldots, x_n)$;
(c) $f_{n+1}(f_n(x_1, x_2, \ldots, x_n), \ldots, f_n(x_1, x_2, \ldots, x_n), x_{n+1}) = f_{n+1}(x_1, x_2, \ldots, x_n, x_{n+1})$.
Find the functions f_n.

Solution. Let's first try to find an expression for $f_1(x), f_2(x, y)$ to help us make an educated guess about the general form of $f_n(x_1, x_2, \ldots, x_n)$.

In (a) we set $n = 1$, $x = 0$ to find $f_1(y) = f_1(0) + y$. From (b) we have that $f_1(0) = 0$. Hence

$$f_1(x) \;=\; x \,.$$

In (a) we now set $n = 2$, $y = -x_1$:

$$f_2(0, x_2 - x_1) \;=\; f_2(x_1, x_2) - x_1 \stackrel{(b)}{\Rightarrow} -f_2(0, x_1 - x_2) \;=\; f_2(x_1, x_2) - x_1 \,.$$

Since f_2 is a symmetric function

$$-f_2(x_1 - x_2, 0) \;=\; f_2(x_1, x_2) - x_1 \,.$$

Again in (a), we set $n = 2$, $y = -x_2$:

$$f_2(x_1 - x_2, 0) \;=\; f_2(x_1, x_2) - x_2 \,.$$

Adding the last two equations, we find

$$f_2(x_1, x_2) \;=\; \frac{x_1 + x_2}{2} \,.$$

The previous results point out that

$$f_n(x_1, x_2, \ldots, x_n) \;=\; \frac{x_1 + x_2 + \cdots + x_n}{n} \,.$$

We shall prove this formula using the method of induction.

For $n = 1, 2$ we have already shown it to be true. We will assume that it is true for $n = k$

$$f_k(x_1, x_2, \ldots, x_k) \;=\; \frac{x_1 + x_2 + \cdots + x_k}{k} \,, \tag{9.2}$$

and we shall show that this implies that the formula is true for $n = k + 1$.

Let $\overline{x}$ stand for the arithmetic mean of $x_1, x_2, \ldots, x_{k+1}$:

$$\overline{x} \;=\; \frac{x_1 + x_2 + \cdots + x_{k+1}}{k + 1} \,.$$

We define the auxiliary variables $a_1, a_2, \ldots, a_{k+1}$ by $a_i = x_i - \overline{x}$. Obviously $\overline{a} = 0$. Equivalently $a_1 + \cdots + a_k + a_{k+1} = 0$ or $A = -a_{k+1}$, where we defined $A = a_1 + \cdots + a_k$. From condition (a) we have that

$$f_{k+1}(x_1, x_2, \ldots, x_{k+1}) \;=\; f_{k+1}(a_1 + \overline{x}, a_2 + \overline{x}, \ldots, a_{k+1} + \overline{x}) \;=\; f_{k+1}(a_1, a_2, \ldots, a_{k+1}) + \overline{x} \,.$$

The proof will be complete if we show that

$$f_{k+1}(a_1, a_2, \ldots, a_{k+1}) \;=\; 0 \,.$$

To this goal, notice that

$$\begin{aligned} f_{k+1}(a_1, a_2, \ldots, a_{k+1}) &= f_{k+1}(f_k(a_1, a_2, \ldots, a_k), \ldots, f_k(a_1, a_2, \ldots, a_k), a_{k+1}) \\ &= f_{k+1}(-\frac{a_{k+1}}{k}, \ldots, -\frac{a_{k+1}}{k}, a_{k+1}) \\ &\overset{(9.2)}{=} f_{k+1}(f_k(-a_{k+1}, 0, \ldots, 0), \ldots, f_k(-a_{k+1}, 0, \ldots, 0), a_{k+1}) \\ &\overset{(c)}{=} f_{k+1}(-a_{k+1}, 0, \ldots, 0, a_{k+1}) \, . \end{aligned}$$

Since the function f_n is symmetric

$$\begin{aligned} f_{k+1}(a_1, a_2, \ldots, a_{k+1}) &= f_{k+1}(a_{k+1}, 0, \ldots, 0, -a_{k+1}) \\ &\overset{(b)}{=} -f_{k+1}(-a_{k+1}, 0, \ldots, 0, a_{k+1}) \, . \end{aligned}$$

In other words

$$f_{k+1}(a_1, a_2, \ldots, a_{k+1}) = -f_{k+1}(a_1, a_2, \ldots, a_{k+1}) \, ,$$

which implies that $f_{k+1}(a_1, a_2, \ldots, a_{k+1}) = 0$ as sought. □

Comment. In the actual Putnam problem given to contestants, the expression of f_n was also given. □

Problem 9.6. *Let $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{m \times m}$ be continuous functions that satisfy the functional relation*

$$\boldsymbol{A}(\vec{x}, \vec{z}) = \boldsymbol{B}(\vec{x}, \vec{y}) \cdot \boldsymbol{C}(\vec{y}, \vec{z}) \, , \tag{9.3}$$

where $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ are nonsingular square matrices. Find all functions $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ that satisfy the above conditions.

Solution. To begin we set $\vec{y} = \vec{a}$ in equation (9.3) where $\vec{a}$ is some constant vector. So we have, after setting $\boldsymbol{B}(\vec{x}, \vec{a}) = \boldsymbol{M}(\vec{x})$ and $\boldsymbol{C}(\vec{a}, \vec{x}) = \boldsymbol{N}(\vec{x})$:

$$\begin{aligned} \boldsymbol{A}(\vec{x}, \vec{z}) &= \boldsymbol{M}(\vec{x}) \cdot \boldsymbol{N}(\vec{z}) \, , \\ \boldsymbol{B}(\vec{x}, \vec{y}) \cdot \boldsymbol{C}(\vec{y}, \vec{z}) &= \boldsymbol{M}(\vec{x}) \cdot \boldsymbol{N}(\vec{z}) \, . \end{aligned}$$

Now if we set $\vec{x} = \vec{b}$ and multiply the second equation on the left by $\boldsymbol{B}^{-1}(\vec{b}, \vec{y})$ we get:

$$\boldsymbol{C}(\vec{y}, \vec{z}) = \boldsymbol{B}^{-1}(\vec{b}, \vec{y}) \cdot \boldsymbol{M}(\vec{b}) \cdot \boldsymbol{N}(\vec{z}) = \boldsymbol{L}(\vec{y}) \cdot \boldsymbol{N}(\vec{z}) \, .$$

If we then plug this into the previous equation and multiply on the right by $\boldsymbol{N}^{-1}(\vec{z}) \cdot \boldsymbol{L}^{-1}(\vec{y})$ we get:

$$\boldsymbol{B}(\vec{x}, \vec{y}) = \boldsymbol{M}(\vec{x}) \cdot \boldsymbol{L}^{-1}(\vec{y}) \, .$$

Therefore, for any non-singular matrices $\boldsymbol{L}(\vec{y})$, $\boldsymbol{M}(\vec{x})$ and $\boldsymbol{N}(\vec{z})$, the solution is:

$$\begin{aligned}
\boldsymbol{A}(\vec{x},\vec{z}) &= \boldsymbol{M}(\vec{x})\cdot\boldsymbol{N}(\vec{z})\,, \\
\boldsymbol{B}(\vec{x},\vec{y}) &= \boldsymbol{M}(\vec{x})\cdot\boldsymbol{L}^{-1}(\vec{y})\,, \\
\boldsymbol{C}(\vec{y},\vec{z}) &= \boldsymbol{L}(\vec{y})\cdot\boldsymbol{N}(\vec{z})\,.
\end{aligned}$$ □

Problem 9.7 (IMO 1981)**.** *The function $f:\mathbb{N}\times\mathbb{N}\to\mathbb{R}$ satisfies the following conditions:*
(a) $f(0,y)=y+1$;
(b) $f(x+1,0)=f(x,1)$;
(c) $f(x+1,y+1)=f(x,f(x+1,y))$.
Find the value $f(4,1981)$.

Solution. We will compute $f(1,y)$, $f(2,y)$, $f(3,y)$, $f(4,y)$, $\forall y$ successively.

We notice

$$\begin{aligned}
f(1,0) &\overset{(b)}{=} f(0,1) \overset{(a)}{=} 2\,, \\
f(1,1) &\overset{(c)}{=} f(0,f(1,0)) = f(0,2) \overset{(a)}{=} 3\,, \\
f(1,2) &\overset{(c)}{=} f(0,f(1,1)) = f(0,3) \overset{(a)}{=} 4\,.
\end{aligned}$$

We therefore suspect that

$$f(1,y) = y+2\,.$$

We shall prove it using the method of induction. Let it be true for $y=n$:

$$f(1,n) = n+2\,.$$

Then, using conditions (a) and (c), it is true for $y=n+1$:

$$f(1,n+1) \overset{(c)}{=} f(0,f(1,n)) = f(0,n+2) \overset{(a)}{=} n+3 = (n+1)+2\,.$$

Now we compute $f(2,y)$. We have:

$$f(2,y) \overset{(c)}{=} f(1,f(2,y-1)) = f(2,y-1)+2\,.$$

The values $f(2,y)$ form an arithmetic progression with step 2. Hence

$$f(2,y) = 2y + f(2,0)\,.$$

Also

$$f(2,0) \overset{(b)}{=} f(1,1) = 3\,.$$

So finally

$$f(2, y) \;=\; 2y + 3\,.$$

Then we compute $f(3, y)$. We have:

$$f(3, y) \stackrel{(c)}{=} f(2, f(3, y-1)) \;=\; 2f(3, y-1) + 3\,.$$

The values $f(3, y)$ form a mixed progression with step 3 and ratio 2. Hence

$$f(3, y) \;=\; 2^y f(3, 0) + 3\,\frac{2^y - 1}{2 - 1}\,.$$

Also

$$f(3, 0) \stackrel{(b)}{=} f(2, 1) \;=\; 5\,,$$

and

$$f(3, y) \;=\; 2^{y+3} - 3\,.$$

Finally we compute $f(4, y)$. We have:

$$f(4, y) \stackrel{(c)}{=} f(3, f(4, y-1)) \;=\; 2^{f(4, y-1)+3} - 3\,.$$

This equation is exactly the recursion relation solved in Problem 3.5. Therefore

$$f(4, y) \;=\; 2^{2^{\cdot^{\cdot^{2^2}}}} - 3\,,$$

where the number 2 appears $y + 3$ times. In particular,

$$f(4, 1981) \;=\; 2^{2^{\cdot^{\cdot^{2^2}}}} - 3\,,$$

where the number 2 appears 1984 times. □

Chapter 10

Systems of Equations

Another 'generalization' one might implement is to search for solutions in a system of functional equations. In this case, of course, the admitted solutions are those in the intersection of the individual sets of solutions. So, in principle one may seek solutions for each equation in the system and then select the common ones.

Problem 10.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function that satisfies the functional equations (5.1) and (5.20). Prove that $f(x) \equiv 0$ or $f(x) = x$.*

Solution. The set of solutions of equation (5.1) is $\{f(x) = a\,x,\ a \in \mathbb{R}\}$. The set of solutions of equation (5.20) is $\{f(x) \equiv 0,\ f(x) = x^b,\ b \in \mathbb{R}\}$. The intersection of the two sets is the set $\{f(x) \equiv 0,\ f(x) = x\}$. ◻

However, given a system of equations, if one uses information from all equations, the solution may be obtained more easily. As an example, let's solve again the previous problem without relying on the solutions of (5.1) and (5.20).

Problem 10.2. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function (not necessarily continuous) that satisfies the Cauchy functional equations (5.1) and (5.20). Prove that $f(x) \equiv 0$ or $f(x) = x$.*

Solution. As we have seen, (5.1) implies that

$$f(q) = f(1)\,q\,, \quad \forall q \in \mathbb{Q}\,.$$

At the same time (5.20) implies that $f(x) \equiv 0$ or $f(x) > 0$, $\forall x \in \mathbb{R}$. In the latter case, $f(1) = 1$. Therefore, either $f(x) \equiv 0$ or

$$f(q) = q\,, \quad \forall q \in \mathbb{Q}\,.$$

Now let $x_1 > x_2$ and set $h = x_1 - x_2$. Since $f(x) > 0$, equation (5.1) gives that

$$f(x_1) = f(x_2 + h) = f(x_2) + f(h) > f(x_2)\,.$$

That is, the function f is strictly increasing.

Now, let's assume that $f(x) \not\equiv x$. Then, there must exist a $x_0 \in \mathbb{R} \setminus \mathbb{Q}$, such that $f(x_0) \neq x_0$. This means that either $f(x_0) > x_0$ or $f(x_0) < x_0$.

If $f(x_0) > x_0$, then we can pick a rational number q_0 such that in the interval $(x_0, f(x_0))$:

$$x_0 < q_0 \;=\; f(q_0) < f(x_0)\,.$$

But this would contradict the fact that f is strictly increasing.

If $f(x_0) < x_0$, then we can pick a rational number q_0' such that in the interval $(f(x_0), x_0)$:

$$f(x_0) < f(q_0) \;=\; q_0 < x_0\,,$$

which would also contradict the fact that f is strictly increasing.

Therefore, it must be that $f(x) \equiv x$. □

Incidentally, let's make a comment motivated by the previous solution: The reader should not underestimate the result proved above. He should notice that once the monotonicity of the function $f(x)$ was established, the functional equations were bypassed in establishing the final result. Given the function $f(q) = q$ defined on $\mathbb{Q}$, there was only one monotonic extension of the function f on $\mathbb{R}$.

Question. *Let* $f : \mathbb{Q} \to \mathbb{R}$ *such that*

$$f(q) \;=\; \begin{cases} -1\,, & \text{if } q < -\sqrt{2}\,, \\ 0\,, & \text{if } -\sqrt{2} < q < +\sqrt{2}\,, \\ +1\,, & \text{if } q > +\sqrt{2}\,, \end{cases}$$

This function is non-decreasing (and thus monotonic). However, there are an infinite number of extensions for f *defined on* $\mathbb{R}$ *depending on the values chosen for* $f(\pm\sqrt{2})$. *Why doesn't the above comment apply here?*

Answer. If one tries to write down a formal proof for a unique extension, then he discovers the following theorem: *Given two continuous functions* f *and* g *defined on* $\mathbb{R}$ *such that* $f(q) = g(q)$ *for all* $q \in \mathbb{Q}$, *then* $f(x) = g(x)$ *for all* $x \in \mathbb{R}$. Notice that the theorem requires continuity. However, the function f given above has no monotonic extension that is continuous at $f(\pm\sqrt{2})$.

Incidentally, we should point out that the theorem can be extended beyond $\mathbb{Q}$ and $\mathbb{R}$: *Given two continuous functions* f *and* g *defined on the set* A *such that* $f(q) = g(q)$ *for all* $q \in S$ *where* S *is a dense subset of* A, *then* $f(x) = g(x)$ *for all* $x \in A$. In other words, if a continuous function is known in a set S and all points of a larger set A are limits of sequences exclusively in S, then the function is uniquely extended in A.

In the light of the above theorem, we may revisit the $\mathbb{NQR}$-method and the proof of the D'Alembert-Poisson I equation (7.21). The $\mathbb{NQR}$-method relies on the fact that $\mathbb{Q}$ is dense in $\mathbb{R}$. So, establishing the function over $\mathbb{Q}$, we can then extend it to $\mathbb{R}$ easily and uniquely. The numbers of the set $D = \{\frac{m}{2^n},\ m, n \in \mathbb{N}\}$ are known as **dyadic numbers**. This set D is known to be dense in $\mathbb{R}$. A continuous or monotonic function in D has a unique extension in $\mathbb{R}$.

10.1 Solved Problems

Problem 10.3. *Find the continuous solutions $F, H, G, g, h, f : \mathbb{R} \to \mathbb{R}$ to the functional equation*

$$F(xy) \;=\; H(x)^{g(y)}\, G(y)^{h(x)} \,,$$

given

$$f(xy) \;=\; h(x)\, g(y) \,.$$

Solution. Using the second, we can rewrite the first functional equation in the form

$$F(xy)^{\frac{1}{g(y)h(x)}} \;=\; F(xy)^{\frac{1}{f(xy)}} \;=\; H(x)^{\frac{1}{h(x)}}\, G(y)^{\frac{1}{g(y)}} \,.$$

This equation can be solved using the Pexider equation (8.11). In other words, if

$$\begin{aligned} f(x) &= a\, b\, x^{\mu} \,, \\ g(x) &= a\, x^{\mu} \,, \\ h(x) &= b\, x^{\mu} \,, \end{aligned}$$

the original equation is equivalent to

$$A(x)\, B(y) \;=\; C(x\, y) \,,$$

where

$$A(x) \;=\; H(x)^{\frac{1}{h(x)}} \,, \quad B(x) \;=\; G(x)^{\frac{1}{g(x)}} \,, \quad C(x) \;=\; F(x)^{\frac{1}{f(x)}} \,.$$

Therefore, $A(x), B(x), C(x)$ also satisfy (8.11) and they must be given by

$$\begin{aligned} C(x) &= A\, B\, x^{\nu} \,, \\ A(x) &= A\, x^{\nu} \,, \\ B(x) &= B\, x^{\nu} \,. \end{aligned}$$

From these expression we can find the functions, $F(x), H(x), G(x)$. Thus the system of equations is solved by the functions

$$\begin{aligned} f(x) &= a\, b\, x^{\mu} \,, & g(x) &= a\, x^{\mu} \,, & h(x) &= b\, x^{\mu} \,, \\ F(x) &= (A\, B\, x^{\nu})^{abx^{\mu}} \,, & G(x) &= (A\, x^{\nu})^{ax^{\mu}} \,, & H(x) &= (B\, x^{\nu})^{bx^{\mu}} \,. \end{aligned}$$

□

Problem 10.4. *Find the continuous solutions $F, H, G, g, h, f : \mathbb{R} \to \mathbb{R}$ to the functional equation*

$$F(x\, y) \;=\; g(y)\, H(x) + h(x)\, G(y)$$

given

$$f(xy) \;=\; h(x)\, g(y) \,.$$

Solution. We can reduce its solution to the preceding problem by exponentiating it:

$$e^{F(x\,y)} \;=\; (e^{H(x)})^{g(y)}\,(e^{G(y)})^{h(x)}\,.$$

So the functional equation is solved by the functions

$$\begin{aligned} f(x) &= a\,b\,x^{\mu}\,, & g(x) &= a\,x^{\mu}\,, & h(x) &= b\,x^{\mu}\,, \\ F(x) &= ab\,x^{\mu}\,\ln(A\,B\,x^{\nu})\,, & G(x) &= ab\,x^{\mu}\,\ln(A\,x^{\nu})\,, & H(x) &= ab\,x^{\mu}\,\ln(B\,x^{\nu})\,. \end{aligned}$$ □

Problem 10.5. *Find the continuous solutions $F, H, G, g, h, f : \mathbb{R} \to \mathbb{R}$ of the following systems of functional equations:*

(a) System 1:

$$\begin{aligned} F(x+y) &= H(x)^{g(y)}\,G(y)^{h(x)}\,, \\ f(x+y) &= h(x)\,g(y)\,, \end{aligned}$$

(b) System 2:

$$\begin{aligned} F(x+y) &= g(y)\,H(x) + h(x)\,G(y)\,, \\ f(x+y) &= h(x)\,g(y)\,. \end{aligned}$$

Solution. The solution proceeds as in the previous two problems. We then arrive at the following solutions:

(a) The System 1 is solved by the functions

$$\begin{aligned} f(x) &= a\,b\,c^{x}\,, \quad g(x) = a\,c^{x}\,, \quad h(x) = b\,c^{x}\,, \\ F(x) &= (A\,B\,d^{x})^{abc^{x}}\,, \quad G(x) = (A\,d^{x})^{abc^{x}}\,, \quad H(x) = (B\,d^{x})^{abc^{x}}\,. \end{aligned}$$

(b) The System 2 is solved by the functions

$$\begin{aligned} f(x) &= a\,b\,c^{x}\,, & g(x) &= a\,c^{x}\,, & h(x) &= b\,c^{x}\,, \\ F(x) &= ab\,c^{x}\,\ln(A\,B\,d^{x})\,, & G(x) &= a\,c^{x}\,\ln(A\,d^{x})\,, & H(x) &= b\,c^{x}\,\ln(B\,d^{x})\,. \end{aligned}$$ □

Problem 10.6. *Find all continuous solutions $f : \mathbb{R}^2 \to \mathbb{R}$ of the system*

$$\begin{aligned} f(x_1+x_2, y) &= f(x_1, y) + f(x_2, y)\,, \quad \forall x_1, x_2, y \in \mathbb{R}\,, \\ f(x, y_1+y_2) &= f(x, y_1) + f(x, y_2)\,, \quad \forall x, y_1, y_2 \in \mathbb{R}\,. \end{aligned}$$

Solution. The first of two given functional equations is just the linear Cauchy equation of one variable — the first variable x of $f(x, y)$. All continuous solutions of this equation are of the form

$$f(x, y) = c(y)\, x\,,$$

where the constant here depends on (the parameter) y. Inserting this result in the second of the given equations, we find:

$$c(y_1 + y_2) = c(y_1) + c(y_2)\,, \quad \forall y_1, y_2 \in \mathbb{R}\,.$$

This is also the linear Cauchy equation with solution

$$c(y) = c\, y\,.$$

Therefore, the allowed continuous solutions to the given system are the functions $f(x, y) = c\, xy$, with c a constant. □

Comment. Please notice the difference between the following two different results: the result of this problem for the continuous function $f(x, y)$ satisfying the *two* functional equations given and the result of Problem 9.1 for the continuous function $F(x, y)$ satisfying a *single* functional equation

$$F(x_1 + x_2, y_1 + y_2) = f(x_1, y_1) + f(x_2, y_2)\,, \quad \forall x_1, x_2, y_1, y_2 \in \mathbb{R}\,.$$

The two results are quite different: $f(x, y) = c\, xy$ in the present case while $F(x, y) = c_1\, x + c_2\, y$ in the case of Problem 9.1. □

Part IV

CHANGING THE RULES

Chapter 11

Less Than Continuity

11.1 Imposing Weaker Conditions

Darboux was the first to show that the Cauchy functional equations can be solved by the same functions even if the condition of continuity was substituted with other milder conditions. The four problems that follow discuss the linear Cauchy functional equation (5.1) with weaker conditions.

Problem 11.1. *Find all functions $f : \mathbb{R} \to \mathbb{R}$ that satisfy the following conditions:*
(a) $f(x) + f(y) = f(x+y)$, $\forall x, y \in \mathbb{R}$, and
(b) f is continuous at a point $x_0 \in \mathbb{R}$.

Solution. We claim that condition (b) implies that f is continuous at every point $x \in \mathbb{R}$, i.e. $\lim_{y \to x} f(y) = f(x)$. To see this, let's set $h = x_0 + y - x$. As $y \to x$, $h \to x_0$ and

$$\begin{aligned}
\lim_{y \to x} f(y) &= \lim_{y \to x} f\left(h + (x - x_0)\right) \\
&= \lim_{y \to x} \left(f(h) + f(x - x_0)\right) \\
&= \lim_{h \to x_0} f(h) + f(x - x_0) \\
&= f(x_0) + f(x - x_0) \; = \; f(x_0 + x - x_0) \; = \; f(x) \, .
\end{aligned}$$

The problem is thus equivalent that of section 5.1 with solution $f(x) = c\,x$, where c is an arbitrary constant. □

Problem 11.2. *Find all functions $f : \mathbb{R} \to \mathbb{R}$ that satisfy the following conditions:*
(a) $f(x) + f(y) = f(x+y)$, $\forall x, y \in \mathbb{R}$, and
(b) f is monotonically increasing.

Solution. As we have seen in section 5.1, condition (a) implies that $f(q) = c\,q$, $\forall q \in \mathbb{Q}$. Now for any $x \in \mathbb{R}$, we can always find an increasing sequence $\{a_n\}_{n\in\mathbb{N}}$ of rational numbers and a decreasing sequence $\{b_n\}_{n\in\mathbb{N}}$ of rational numbers which converge to x. Then

$$a_n \le x \le b_n\,, \quad \forall n \in \mathbb{N}\,.$$

From condition (b), it follows that

$$f(a_n) \le f(x) \le f(b_n) \Rightarrow ca_n \le f(x) \le cb_n\,, \quad \forall n \in \mathbb{N}\,.$$

By taking the limit $n \to \infty$, we find $f(x) = c\,x$. □

Problem 11.3. *Find all functions $f : \mathbb{R} \to \mathbb{R}$ that satisfy the following conditions:*
(a) $f(x) + f(y) = f(x + y)$, $\forall x, y \in \mathbb{R}$, and
(b) there exists a positive number ε such that $f(x) \ge 0$, $\forall x \in (0, \varepsilon)$.

Solution. We will show that condition (b) implies that the function is monotonically increasing, and therefore the problem is equivalent to the previous one.

Let $x_1 > x_2$ such that $h = x_1 - x_2 \in (0, \varepsilon)$. Then $f(h) \ge 0$ and

$$f(x_1) = f(x_2 + h) = f(x_2) + f(h) \ge f(x_2) + 0 = f(x_2)\,.$$

Now, notice that by using this relation in the intervals $[n\varepsilon/2, (n + 1)\varepsilon/2]$, $n \in \mathbb{Z}$, we can show straightforwardly that f is non-decreasing. □

Problem 11.4. *Find all continuous functions $f : \mathbb{R} \to \mathbb{R}$ which satisfy the following conditions:*
(a) $f(x + y) = f(x) + f(y)$, $\forall x, y \in \mathbb{R}$;
(b) there exists a positive number ε such that f is bounded on an interval $(0, \varepsilon)$.

Solution. Condition (a) implies that

$$f(q\,x) = q\,x\,, \quad \forall q \in \mathbb{Q}\,, \quad \forall x \in \mathbb{R}\,.$$

(See equation (5.7).)

Now, let $\{x_n\}$ be an arbitrary sequence of positive real numbers such that $\lim\limits_{n\to\infty} x_n = 0$. Given $\{x_n\}$, we select a sequence of positive rational numbers $\{q_n\}$ as follows. For any n, the rational number q_n is chosen such that

$$\frac{1}{\sqrt{x_n}} < q_n < \frac{1}{\sqrt[3]{x_n}}\,.$$

Then, in the limit $n \to \infty$, $q_n \to +\infty$. At the same time

$$\sqrt{x_n} < q_n\, x_n < \sqrt[3]{x_n^2}\,,$$

i.e. $\lim\limits_{n\to+\infty}(q_n\, x_n) = 0$. Since the sequence $(q_n\, x_n)$ converges to zero, eventually all terms of the sequence will fall in the interval $(0, \varepsilon)$. That is, there exists $n_0 \in \mathbb{N}$ such that $q_n\, x_n \in (0, \varepsilon)$, $\forall n > n_0$.

Condition (b) implies that there exists $M \in \mathbb{R}_+$ such that $|f(x)| \leq M$, $\forall x \in (0, \varepsilon)$. Therefore $|f(q_n\, x_n)| \leq M$, $\forall n > n_0$. Then

$$|f(x_n)| \;=\; \left| f\left(\frac{1}{q_n}\, q_n x_n\right)\right| \;=\; \frac{1}{q_n}\,|f(q_n\, x_n)| \leq \frac{M}{q_n} \to 0\,, \quad \text{when} \quad n \to \infty\,.$$

In other words, for any arbitrary sequence $\{x_n\}$ which converges to zero, $\lim\limits_{n\to+\infty} f(x_n) = 0 = f(0)$ and therefore the function is continuous at $x = 0$. As we have proved above, continuity at one point gives $f(x) = c\, x$, $\forall x \in \mathbb{R}$. □

11.2 Non-Continuous Solutions

Because of the results in the previous section, mathematicians initially thought that the solution to the linear Cauchy equation (5.1) had always to be $f(x) = c\, x$. However, Hamel in 1905, using ideas from set theory, proved [47] that this is not the case. His solution makes use of the now so-called *Hamel basis*. In this section we shall explain this result. Unfortunately, since no known explicit example of a Hamel basis exists, we will first use an artificial example in order to get some intuition for what follows.

An Analogy

Let's consider the set

$$A \;=\; \mathbb{Q} + \mathbb{Q}\sqrt{2} + \mathbb{Q}\sqrt{3} \;=\; \{r + s\sqrt{2} + t\sqrt{3} \mid r, s, t \in \mathbb{Q}\}\,.$$

Every number in A is uniquely expressed as a linear combination of the elements of the set

$$H \;=\; \{1,\ \sqrt{2},\ \sqrt{3}\}\,.$$

The set H consists of the 'Hamel basis' of A.

Now let's solve the linear Cauchy equation (5.1) for functions $f : A \to A$. To this end, we can repeat the method of Section 5.1. Without writing down all the details, we show that

$$\begin{gathered} f((r_1 + r_2 + \cdots + r_n) + (s_1 + s_2 + \cdots + s_n)\sqrt{2} + (t_1 + t_2 + \cdots + t_n)\sqrt{3}) \;= \\ f((r_1 + r_2 + \cdots + r_n)) + f((s_1 + s_2 + \cdots + s_n)\sqrt{2}) + f((t_1 + t_2 + \cdots + t_n)\sqrt{3})\,, \end{gathered}$$

from which we eventually arrive at

$$f(r + s\sqrt{2} + t\sqrt{3}) \;=\; r\,f(1) + s\,f(\sqrt{2}) + t\,f(\sqrt{3})\,.$$

Now let's re-express this result using the 'Hamel basis'. Let $\tilde{f}$ be any mapping from H to A, $\tilde{f} : H \to A$. Then the function f which assigns to each $x = r + s\sqrt{2} + t\sqrt{3}$ the value

$$f(x) \;=\; r\,\tilde{f}(1) + s\,\tilde{f}(\sqrt{2}) + t\,\tilde{f}(\sqrt{3})$$

is a solution of the linear Cauchy functional equation in A. Given all possible mappings $\tilde{f}$ of H, we get all possible solutions of the functional equation.

An example of $\tilde{f} : H \to A$ is

$$\tilde{f}(1) \;=\; q_0\,, \quad \tilde{f}(\sqrt{2}) \;=\; q_0\sqrt{2}\,, \quad \tilde{f}(\sqrt{3}) \;=\; q_0\sqrt{3}\,,$$

for some fixed $q_0 \in \mathbb{Q}$. Then

$$f(x) \;=\; q_0\,x\,.$$

This is a linear function with $f(x) = 0$ iff $x = 0$.

Another example of $\tilde{f} : H \to A$ is

$$\tilde{f}(1) \;=\; 1\,, \quad \tilde{f}(\sqrt{2}) \;=\; \tilde{f}(\sqrt{3}) \;=\; 0\,.$$

Then

$$f(r + s\sqrt{2} + t\sqrt{3}) \;=\; r\,.$$

Notice that this is not linear function. In particular, it vanishes for an infinite number of points: $f(x) = 0$ if $x = 0$ or $x = s\sqrt{2} + t\sqrt{3}$.

Hamel's Solution

Definition 11.1. A non-denumerable subset H of $\mathbb{R}$ is called a **Hamel basis**, if for every real number x there exists a finite natural number n and elements $h_1, h_2, \ldots, h_n$ of H such that x is written as a sum

$$x \;=\; \sum_{i=1}^{n} q_i\,h_i\,,$$

with rational coefficients q_i. We shall call this expression the **Hamel decomposition** of x in H.

Strictly speaking, the definition of a Hamel basis does not guarantee its existence. Hamel's major contribution to the problem was to prove, using the axiom of choice, the following theorem.

Theorem 11.1 (Hamel)**.** *At least one Hamel basis exists.*

Since set theory is not the main focus of these introductory notes, we shall omit the proof. We are now in position to solve Cauchy's equation without any constraint on the functions f.

Theorem 11.2. *The most general solution to the linear Cauchy equation (5.1) is found by*
(i) introducing a Hamel basis H in $\mathbb{R}$,
(ii) assigning arbitrary values for f at the points of the Hamel basis, and
(iii) if $x \in \mathbb{R}$ has a Hamel decomposition $x = \sum_i q_i h_i$, then $f(x) = \sum_i q_i f(h_i)$.

The proof of this theorem is straightforward.

Proof. (Direct proposition) Given (5.1) one can show that

$$f\left(\sum_{i=1}^{n} y_i\right) = \sum_{i=1}^{n} f(y_i)\,, \quad \forall n \in \mathbb{N}^*, \quad \forall y_i \in \mathbb{R}\,, i = 1, 2, \ldots, n\,,$$
$$f(q\,y) = q\,f(y)\,, \quad \forall y \in \mathbb{R}\,, \quad \forall q \in \mathbb{Q}\,.$$

From these and the Hamel decomposition of any number $x = \sum_i q_i\, h_i$, we find

$$f(x) = f\left(\sum_i q_i\, h_i\right) = \sum_i f(q_i\, h_i) = \sum_i q_i\, f(h_i)\,.$$

(Inverse proposition) Let $x, x' \in \mathbb{R}$ with Hamel decompositions

$$x = \sum_{\{h_i\}} q_i h_i\,, \quad x' = \sum_{\{h'_j\}} q'_j h'_j\,,$$

where $\{h_i\}$, $\{h'_j\}$ stand for the set of Hamel elements in the decomposition of x, x' respectively. The two sets do not necessarily have the same number of elements nor do they necessarily include the same elements of H. Let $\{\tilde{h}_k\} = \{h_i\} \cup \{h'_j\}$. By adding a convenient number of terms with zero coefficients

$$x = \sum_{\{\tilde{h}_k\}} \tilde{q}_k \tilde{h}_k\,, \quad x' = \sum_{\{\tilde{h}_k\}} \tilde{q}'_k \tilde{h}_k\,.$$

Then

$$f(x) = \sum_{\{\tilde{h}_k\}} f(\tilde{h}_k)\,\tilde{q}_k\,, \quad f(x') = \sum_{\{\tilde{h}_k\}} f(\tilde{h}_k)\,\tilde{q}'_k\,,$$

and

$$f(x) + f(x') = \sum_{\{\tilde{h}_k\}} (\tilde{q}_k + \tilde{q}'_k)\, f(\tilde{h}_k)\,.$$

On the other hand,

$$x + x' = \sum_{\{\tilde{h}_k\}} (\tilde{q}_k + \tilde{q}'_k)\, \tilde{h}_k\,,$$

and

$$f(x+x') = \sum_{\{\tilde{h}_k\}} (\tilde{q}_k + \tilde{q}'_k)\, f(\tilde{h}_k)\,,$$

which shows that

$$f(x+x') = f(x) + f(x')\,. \qquad \square$$

Corollary. *The Hamel solution of (5.1) is continuous if and only if there exists a real number c such that $f(h) = c\,h$, for all elements h of the Hamel basis.*

Example 11.1. To understand Hamel's solution, let's construct one illustrative solution. Let h_0 be any element of the Hamel basis H and define $f(h_0) = 1$ and $f(h) = 0$, $\forall h \in H \setminus \{h_0\}$. Then the function

$$f(x) = \begin{cases} q_0\,, & \text{if } x = q_0 h_0 + \ldots\,, \\ 0\,, & \text{otherwise}\,, \end{cases}$$

is a discontinuous solution of (5.1). $\square$

With very little additional effort, we can find the most general solutions to the rest of the Cauchy equations.

Problem 11.5. *Give the most general solution to the exponential Cauchy functional equation (5.8). Then show how we can recover the solution $f(x) = a^x$ from the general one.*

Solution. As seen in Section 5.2, the function $g(x) = \ln f(x)$ satisfies the (5.1) functional equation. Therefore, its most general solution is found by assigning arbitrary values $g(h)$ for the elements of the Hamel basis H, and for every $x \in \mathbb{R}$ with Hamel decomposition $x = \sum_i q_i h_i$, we write

$$g(x) = \sum_i q_i\, g(h_i)\,.$$

Returning back to function f,

$$\ln f(x) = \sum_i q_i \ln f(h_i) = \sum_i \ln f(h_i) q_i = \ln \prod_i f(h_i)^{q_i} \Rightarrow f(x) = \prod_i f(h_i)^{q_i}\,,$$

where $f(h)$ is arbitrary for $h \in H$.

The function $g(x)$ will be continuous, if there is a number c such that $g(h) = c\,h$, $\forall h \in H$. This implies $\ln f(h) = c\,h$ or $f(h) = e^{ch} = a^h$, $\forall h \in H$, and $a = e^c$. Therefore,

$$f(x) = \prod_i f(h_i)^{q_i} = \prod_i a^{h_i q_i} = a^{\sum_i h_i q_i} = a^x\,. \qquad \square$$

For the other two solutions, we will need the following lemma.

Lemma 11.1. *There exists a non-denumerable set Δ of non-negative elements $\delta \in \mathbb{R}_+^*$ such that any $x \in \mathbb{R}_+^*$ can be written as finite product*

$$x = \prod_{i=1}^{n} \delta_i^{q_i} ,$$

for some $n \in \mathbb{N}^*$ and n rational numbers $q_1, q_2, \ldots, q_n$.

Proof. Let $x \in \mathbb{R}_+^*$. Then there exists a $u \in \mathbb{R}$ such that $x = e^u$. Using a Hamel basis, u can be expressed as a linear sum $u = \sum_{i=1}^n q_i h_i$, for some n and $q_i \in \mathbb{Q}$, $i = 1, 2, \ldots, n$. Then

$$x = e^u = e^{\sum_{i=1}^n q_i h_i} = \prod_{i=1}^{n} (e^{h_i})^{q_i} .$$

Therefore, the set $\Delta = \{\delta = e^h : h \in H\}$ has the required property:

$$x = \prod_{i=1}^{n} \delta_i^{q_i} . \qquad \square$$

Example 11.2. Although no basis is known for $\mathbb{R}_+^*$, we can present a very good example to convey the idea. If instead of $\mathbb{R}_+^*$ we consider $\mathbb{Q}_+^*$, then let Δ be the set of all prime numbers. Then any $x \in \mathbb{Q}_+^*$ can be written as

$$x = \prod_{i=1}^{n} p_i^{k_i} ,$$

for some integer exponents k_i. $\square$

Problem 11.6. *Show that the most general solution to the logarithmic Cauchy equation (5.14) is found by*
(i) introducing a basis Δ in $\mathbb{R}_+$,
(ii) assigning arbitrary values for f at the points of the basis, and
(iii) if $x \in \mathbb{R}_+^*$ has a decomposition $x = \prod_i \delta_i^{q_i}$, then $f(x) = \sum_i q_i f(\delta_i)$.

Solution. (Direct proposition) Given (5.14) one can show that

$$f\left(\prod_{i=1}^{n} y_i\right) = \sum_{i=1}^{n} f(y_i) , \quad \forall n \in \mathbb{N}^*, \quad \forall y_i \in \mathbb{R} , i = 1, 2, \ldots, n ,$$
$$f(y^q) = q\, f(y_i) , \quad \forall y \in \mathbb{R} , \quad \forall q \in \mathbb{Q} .$$

From these and the decomposition of any positive number $x = \prod_i \delta_i^{q_i}$, we find

$$f(x) = f\left(\prod_i \delta_i^{q_i}\right) = \sum_i f(\delta_i^{q_i}) = \sum_i q_i\, f(\delta_i) .$$

(Inverse proposition) Let $x, x' \in \mathbb{R}_+^*$ with decompositions

$$x \;=\; \prod_i \delta_i^{q_i}\,, \qquad x' \;=\; \prod_j (\delta_j')^{q_j'}\,,$$

where $\{\delta_i\}$, $\{\delta_j'\}$ stand for the set of Hamel elements in the decomposition of x, x' respectively. The two sets do not necessarily have the same number of elements and nor do they necessarily include the same elements of Δ. Let $\{\tilde{\delta}_k\} = \{\delta_i\} \cup \{\delta_j'\}$. By adding a convenient number of factors with zero exponents

$$x \;=\; \prod_{\{\tilde{\delta}_k\}} \tilde{\delta}_k^{\tilde{q}_k}\,, \qquad x' \;=\; \prod_{\{\tilde{\delta}_k\}} \tilde{\delta}_k^{\tilde{q}_k'}\,,$$

Then

$$f(x) \;=\; \sum_{\{\tilde{\delta}_k\}} f(\tilde{\delta}_k)\,\tilde{q}_k\,, \qquad f(x') \;=\; \sum_{\{\tilde{\delta}_k\}} f(\tilde{\delta}_k)\,\tilde{q}_k'\,,$$

and

$$f(x) + f(x') \;=\; \sum_{\{\tilde{\delta}_k\}} (\tilde{q}_k + \tilde{q}_k')\, f(\tilde{\delta}_k)\,.$$

On the other hand,

$$x\,x' \;=\; \prod_{\{\tilde{\delta}_k\}} \tilde{\delta}_k^{\tilde{q}_k + \tilde{q}_k'}\,,$$

and

$$f(x\,x') \;=\; \sum_{\{\tilde{h}_k\}} (\tilde{q}_k + \tilde{q}_k')\, f(\tilde{\delta}_k)\,,$$

which shows that

$$f(x\,x') \;=\; f(x) + f(x')\,.$$

The solution of (5.14) is continuous if and only if there exists a real number c such that $f(\delta) = c\,\ln\delta$, for all elements δ of Δ:

$$f(x) \;=\; \sum_i q_i\, f(\delta_i) \;=\; c\sum_i q_i\,\ln\delta_i \;=\; c\,\ln\prod_i \delta_i^{q_i} \;=\; c\,\ln x \;=\; \log_\gamma x\,,$$

where we defined $\gamma = e^{1/c}$. □

Problem 11.7. *Give the most general solutions to the power Cauchy functional equation (5.20).*

Solution. As seen in Section 5.4, the function $g(x) = \ln f(x)$ satisfies the (5.14) functional equation. Therefore, its most general solution is found by assigning arbitrary values $g(\delta)$ for the elements of the basis Δ, and for every $x \in \mathbb{R}$ with decomposition $x = \prod_i \delta_i^{q_i}$, we write

$$g(x) \;=\; \sum_i q_i\, g(\delta_i)\,.$$

Returning back to function f,

$$\ln f(x) = \sum_i q_i \ln f(\delta_i) = \sum_i \ln\left(f(\delta_i)\right)^{q_i} = \ln \prod_i f(\delta_i)^{q_i} \Rightarrow f(x) = \prod_i f(\delta_i)^{q_i} ,$$

where $f(\delta)$ are arbitrary for $\delta \in \Delta$.

The function $g(x)$ will be continuous, if there is a number c such that $g(\delta) = c \ln \delta$, $\forall \delta \in \Delta$. This implies $\ln f(\delta) = c \ln \delta$ or $f(\delta) = \delta^c$, $\forall \delta \in \Delta$. Therefore

$$f(x) = \prod_i f(\delta_i)^{q_i} = \prod_i (\delta_i^{q_i})^c = \left(\prod_i \delta_i^{q_i}\right)^c = x^c . \qquad \square$$

11.3 Solved Problems

Problem 11.8. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function that satisfies*
(a) $f(x + y) = f(\lambda x) + f(\lambda y)$, $\forall x, y \in \mathbb{R}$, with λ some real number, and
(b) f is continuous at $x = 0$.
Find all such functions.

Solution 1. Setting $x = y = 0$ in (a), we find that $f(0) = 0$. We shall examine the following cases:

(i) $\lambda = 1$: Then f satisfies (5.1) and it is continuous at $x = 0$. Therefore $f(x) = c\,x$, where c is a real constant or $f(x) \equiv 0$.

(ii) $\lambda = 0$: Then $f(x + y) = 2f(0) = 0$, i.e. $f(x) \equiv 0$.

(iii) $\lambda = -1$: Setting $y = 0$ in (a), we find $f(x) = f(-x)$. Also, setting $y = -x$ we find $f(-x) = -f(x)$. These two relations imply that $f(x) \equiv 0$.

(iv) $|\lambda| < 1$: Setting $y = 0$ in (a), we find $f(x) = f(\lambda x)$. Applying this equation n times

$$f(x) = f(\lambda^n x) , \quad \forall n \in \mathbb{N}^* , \quad \forall x \in \mathbb{R} .$$

Since $|\lambda| < 1$, $\lim\limits_{n\to\infty} \lambda^n = 0$. Then, by the continuity of f at $x = 0$:

$$0 = f(0) = f\Big(\lim_{n\to\infty} \lambda^n x\Big) = \lim_{n\to\infty} f(\lambda^n x) = \lim_{n\to\infty} f(x) = f(x) .$$

(v) $|\lambda| > 1$: In this case $\mu = 1/\lambda$, with $|\mu| < 1$. Setting $y = 0$, $x = \mu z$ in (a), we find $f(z) = f(\mu z)$. Now, by working exactly as in case (iv), we arrive again at $f(x) \equiv 0$. $\square$

Solution 2. Set x/λ in place of x and μx in place of y, where μ is a real number to be determined shortly:

$$f\left(\left(\frac{1}{\lambda} + \mu\right)x\right) = f(x) + f(\lambda\mu x) .$$

Now we select μ such that

$$\frac{1}{\lambda} + \mu \;=\; \lambda\mu \;\Rightarrow\; \mu \;=\; \frac{1}{\lambda(\lambda-1)}\,.$$

So when $\lambda \neq 0, 1$, the last equation implies $f(x) \equiv 0$.

For $\lambda = 0$, it is immediate that $f(x) \equiv 0$ and for $\lambda = 1$ we recover the continuous solutions of the linear Cauchy equation. □

Solution 3. If $\lambda = 0$, it is immediate that $f(x) \equiv 0$. Defining the function $g(x) = f(\lambda x)$, $\lambda \neq 0$, the functional equation given in (a) becomes the linear Pexider equation:

$$f(x+y) \;=\; g(x) + g(y)\,.$$

The functions f, g are continuous at $x = 0$. This can be used to show that they are continuous at all points as we did with the linear Cauchy equation in Problem 11.1. Indeed, if we set $h = y - x$, then $\lim\limits_{y\to x} h = 0$ and

$$\begin{aligned} \lim_{y\to x} f(y) &= \lim_{y\to x} f\,(h+x) \\ &= \lim_{y\to x} \left(g(h) + g(x)\right) \\ &= \lim_{h\to 0} g(h) + g(x) \\ &= g(0) + g(x) \;=\; f(x)\,. \end{aligned}$$

As we have established, the continuous solutions of the linear Pexider equation are $f(x) = c\,x + 2a$, $g(x) = c\,x + a$, or

$$f(x) \;=\; \frac{c}{\lambda}\,x + a \;=\; c\,x + 2a\,, \quad \forall x \in \mathbb{R}\,.$$

Obviously $a = 0$. If $\lambda = 1$ then $c =$ arbitrary; if $\lambda \neq 0, 1$ then $c = 0$. □

Problem 11.9 (Russia 1997)**.** *For which α does there exist a non-constant function $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(\alpha(x+y)) \;=\; f(x) + f(y)\;?$$

Solution. If $x = y = 0$ then $f(0) = 0$. Then if $\alpha = 0$, we set $y = 0$ to find $f(x) \equiv 0$. If $\alpha \neq 0$, we substitute x/α for x and y/α for y:

$$f(x+y) \;=\; f(\lambda x) + f(\lambda y)\,,$$

where we set $\lambda = 1/\alpha$. From the previous problem we can see that a non-constant solution exists only for the case $\alpha = 1$. (Since continuity or some other condition is not given however, the solutions are not necessarily of the form $f(x) = cx$.) □

Problem 11.10. *Show that if the functions $f, g : \mathbb{R} \to \mathbb{R}$ satisfy the equation*

$$f(\lambda x, \lambda y) = g(\lambda) f(x, y) , \quad \forall \lambda, x, y ,$$

then g must be a solution of the power Cauchy equation.

Solution. Substituting $\lambda\mu$ in place of λ in the given functional equation we find:

$$f(\lambda\mu x, \lambda\mu y) = g(\lambda\mu) f(x, y) .$$

Then we rewrite the left hand side of this equation as follows:

$$\begin{aligned} f(\lambda\mu x, \lambda\mu y) &= f(\lambda(\mu x), \lambda(\mu y)) \\ &= g(\lambda) f(\mu x, \mu y) \\ &= g(\lambda) g(\mu) f(x, y) . \end{aligned}$$

Combining the two equations we conclude that

$$g(\lambda\mu) = g(\lambda) g(\mu) . \qquad \square$$

Comment. If we require that g is continuous, then $g(x) = x^r$ and the given equation reduces to the Euler equation. (See Problem 2.11.) $\square$

Problem 11.11 ([4], Problem 1690)**.** *Prove that there exist functions $f : \mathbb{R} \to \mathbb{R}$ that satisfy*

$$f(x - f(y)) = f(x) + y , \quad \forall x, y \in \mathbb{R} ,$$

and show how such functions can be constructed.

Solution. In the defining equation, we set $x = f(y)$:

$$f(0) = f^2(y) + y . \tag{11.1}$$

This implies that f is one-to-one since

$$f(x) = f(y) \Rightarrow f^2(x) = f^2(y) \Rightarrow f(0) - x = f(0) - y \Rightarrow x = y .$$

Then we set $y = 0$ in the defining equation to find

$$f(x - f(0)) = f(x) \Rightarrow x - f(0) = x \Rightarrow f(0) = 0 .$$

Returning to equation (11.1) we have thus found

$$f^2(y) = -y . \tag{11.2}$$

Finally, in the defining equation we replace y by $f(y)$:

$$f(x - f^2(y)) \;=\; f(x) + f(y)\,,$$

which, with the help of (11.2), reduces to the linear Cauchy equation:

$$f(x + y) \;=\; f(x) + f(y)\,.$$

The function f satisfies the original equation iff it satisfies the linear Cauchy equation and (11.2).

We can construct the solutions using Hamel's basis. Given any real number $x = \sum_i q_i h_i$, since f satisfies the linear Cauchy equation

$$f(x) \;=\; \sum_i q_i\, f(h_i)\,,$$

and

$$f^2(x) \;=\; \sum_i q_i\, f^2(h_i)\,.$$

To satisfy equation (11.2), we only need to consider the restriction of f on H. More precisely, we need to define the values $f(h_i)$ such that $f^2(h_i) = -h_i$. To this end, we partition H into two disjoint sets H_+ and H_- each having the power of continuum. There is a bijective function between H_+ and H_- which we shall use to label the elements of the two sets: to each element h^+_α of H_+, there is a corresponding element h^-_α of H_-. Now we define $f : H \to H$ such that

$$f(h^+_\alpha) \;=\; -h^-_\alpha\,, \quad f(h^-_\alpha) \;=\; +h^+_\alpha\,. \qquad \square$$

Problem 11.12 ([4], Problem 1626)**.** *Let $f, g, h : \mathbb{R} \to \mathbb{R}$ be functions such that*

$$f(g(0)) \;=\; g(f(0)) \;=\; h(f(0)) \;=\; 0\,, \tag{11.3}$$

and

$$f(x + g(y)) \;=\; g(h(f(x))) + y\,, \tag{11.4}$$

for all $x, y \in \mathbb{R}$. Prove that $h = f$ and that

$$g(x + y) \;=\; g(x) + g(y)\,,$$

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

Solution. Setting $x = y = 0$ in the defining equation (11.4), we find

$$f(g(0)) = g(h(f(0))) \Rightarrow g(0) = 0 .$$

Then setting only $x = 0$ in (11.4):

$$f(g(y)) = y . \tag{11.5}$$

In particular, for $y = 0$ this equation implies $f(0) = 0$. Similarly, setting $y = 0$, $x = g(y)$ in (11.4), we find:

$$f(g(y)) = g(h(f(g(y)))) \Rightarrow y = g(h(y)) . \tag{11.6}$$

Using this result to rewrite the first term in the right hand side of the original equation (11.4), we have

$$f(x + g(y)) = f(x) + y , \tag{11.7}$$

This equation produces both results sought. Upon substituting $y = h(z)$

$$f(x + z) = f(x) + h(z) .$$

For $x = 0$ the last equation implies that $f = h$ (and therefore f satisfies the linear Cauchy equation). Also, notice that equations (11.5) and (11.6) now imply that f and g are inverse to each other. If instead, in equation (11.7), we set $x = g(z)$ we find

$$f(g(z) + g(y)) = z + y \Rightarrow g(z) + g(y) = g(z + y) . \qquad \square$$

The following problem given in the 1998 Romanian Olympiad is a duplication of Problem 11.2 but for the Vincze I functional equation (8.13). Instead of continuity, monotonicity is given for one of the unknown functions. Using the $\mathbb{NQR}$-method, one can construct the solution in $\mathbb{Q}$ and then its extension[1] in $\mathbb{R}$.

Problem 11.13 (Romania 1998)**.** *Find all functions $u : \mathbb{R} \to \mathbb{R}$ for which there exists a strictly monotonic function $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(x + y) = f(x)u(y) + f(y) , \quad \forall x, y \in \mathbb{R} .$$

Solution. For $y = 0$, $f(x)(1 - u(0)) = f(0)$. If $u(0) \neq 1$, then $f(x) = f(0)/(1 - u(0))$, that is the function $f(x)$ will be constant and thus not strictly monotonic. Therefore, it must necessarily be $u(0) = 1$. This also implies $f(0) = 0$. If, in the defining equation, we set x in place of y and vice versa, we find

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

[1]To use the statement made in the box of page 146: Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there is a unique monotonic extension of a function from $\mathbb{Q}$ to $\mathbb{R}$.

Therefore

$$f(x)u(y) + f(y) \;=\; f(y)u(x) + f(x)\,.$$

Since, by the monotonicity of $f(x)$ and that $f(0) = 0$, $f(x) \neq 0$ for all $x \neq 0$, we can separate the terms depending on x from the terms depending on y (and thus the name **separation of variables**):

$$\frac{u(x) - 1}{f(x)} \;=\; \frac{u(y) - 1}{f(y)}\,,$$

for all $x, y \neq 0$. Obviously, each term must be a constant, say a:

$$\frac{u(x) - 1}{f(x)} \;=\; a \;\Rightarrow\; u(x) \;=\; af(x) + 1\,.$$

If $a = 0$, $u(x) = 1$ and $f(x + y) = f(x) + f(y)$. The solution is given in Problem 11.2 in page 153: $f(x) = cx$, $x \in \mathbb{R}$. If $a \neq 0$, then

$$\begin{aligned} u(x + y) &= af(x + y) + 1 \;=\; a[f(x)u(y) + f(y)] + 1 \\ &= [af(x) + 1 - 1]u(y) + af(y) + 1 \;=\; [u(x) - 1]u(y) + u(y) \\ &= u(x)u(y)\,. \end{aligned}$$

That is, the function $u(x)$ satisfies the exponential Cauchy equation which is solved by $u(x) = b^x$ for $x \in \mathbb{Q}$. Since $u(x) = af(x) + 1$, if $f(x)$ is strictly monotonic, $u(x)$ is strictly monotonic too. Using the same idea as in Problem 11.2, we conclude that $u(x) = b^x$ and $f(x) = (b^x - 1)/a$ for $x \in \mathbb{R}$. □

The following problem, given by Korea in 1997, is really the same as the previous problem but for the Vincze II functional equation (8.19).

Problem 11.14 (Korea 1997)**.** *Find all pairs of functions $f, u : \mathbb{R} \to \mathbb{R}$ such that*
(a) if $x < y$, then $f(x) < f(y)$;
(b) for all $x, y \in \mathbb{R}$, $f(xy) \;=\; f(x)u(y) + f(y)$.

Solution. For $y = 0$, $f(x)u(0) = 0$. The function $f(x)$ cannot be constant, in particular $f(x) \equiv 0$, since it is strictly increasing; therefore $u(0) = 0$.

If, in the defining equation, we set x in place of y and vice versa, we find

$$f(yx) \;=\; f(y)\,u(x) + f(x)\,.$$

For $y = 0$, it gives $f(0)(1 - u(x)) = f(x)$. $f(0)$ cannot vanish since this would imply that $f(x) \equiv 0$. We define $a = -1/f(0)$ to write

$$u(x) \;=\; af(x) + 1\,.$$

Then

$$
\begin{aligned}
u(xy) &= af(xy)+1 \;=\; a[f(x)\,u(y)+f(y)]+1 \\
&= [af(x)+1-1]\,u(y)+af(y)+1 \;=\; [u(x)-1]\,u(y)+u(y) \\
&= u(x)\,u(y)\,.
\end{aligned}
$$

That is, the function $u(x)$ satisfies the power Cauchy equation for all values of $x \in \mathbb{R}$. Using the $\mathbb{NQR}$-method, we can solve this equation for $x \in \mathbb{Q}_+^*$: $u(x) = x^c$. If f is strictly increasing, then $u(x) = af(x)+1$ is strictly monotonic and thus $u(1) \neq 0$. At the same time, $u(1\cdot 1) = u(1)u(1) \Rightarrow u(1) = 1$ and we conclude that $u(x)$ is strictly increasing (and thus $a > 0$). So finally $u(x) = x^c$ with $c > 0$ for $x \in \mathbb{Q}_+$. To find the values for negative rationals, we observe that $u(-x) < 0 < u(x)$, $x > 0$. Also $u(x^2) = u(-x)^2$ or $u(-x) = -\sqrt{u(x^2)} = -x^c$. Using the same idea as in Problem 11.2, we conclude that

$$
u(x) \;=\; \begin{cases} x^c\,, & \text{if } x \geq 0\,, \\ -|x|^c\,, & \text{if } x < 0\,. \end{cases}
$$

Also $f(x) = (u(x)-1)/a$, with $a > 0$. $\square$

Chapter 12

More Than Continuity

So far we have studied functional equations for functions that are continuous or satisfy weaker conditions. Often, however, the function may be required to satisfy stronger conditions. Probably the most common and most useful condition is differentiability. When a function is differentiable, it is possible to reduce the given functional equation to a differential equation, which can be solved more easily.

12.1 Differentiable Functions

Problem 12.1. *Find all differentiable solutions $f : \mathbb{R} \to \mathbb{R}$ of the linear Cauchy equation (5.1).*

Solution. Differentiating (5.1) with respect to x, we find

$$f'(x+y) \;=\; f'(x)\,.$$

Now we set $x = 0$ and $f'(0) = c$. Then

$$f'(y) \;=\; c \;\Rightarrow\; f(y) \;=\; c\,y + b\,.$$

From (5.1) we can find that $f(0) = 0$ and therefore $b = 0$, giving the final result $f(x) = c\,x$. □

Comment. We notice that continuity or differentiability of $f(x)$ satisfying (5.1) leads to the same set of solutions for this equation. Therefore, one expects to be able to show that *any continuous solution of* (5.1) *must necessarily be differentiable*. Indeed, this can be done as shown below.

Since $f(x)$ is continuous in $\mathbb{R}$ it cannot approach $\pm\infty$ for any $x \in [0,1]$ since, if it did, it would be discontinuous at that point. We thus assume that $f(x)$ is integrable on the interval [0,1]. We then integrate equation (5.1) with respect to y from 0 to 1 to get

$$\int_0^1 f(x+y)\,dy \;=\; f(x) + \int_0^1 f(y)\,dy\,,$$

or

$$f(x) = \int_x^{1+x} f(u)\,du - \int_0^1 f(y)\,dy\,.$$

The second integral in the right hand side is a constant, which we denote by $-C$:

$$f(x) = \int_x^{1+x} f(u)\,du + C\,.$$

We notice now that the right hand side is a differentiable function; the left hand side must therefore be differentiable, too. In particular,

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

The observation that continuity implies differentiability for (at least some) functional equations is important since it allows to convert functionals equations to differential equations — a subject which allows systematic treatment of the problems posed. (For additional information the reader should consult Section 12.3.) □

Problem 12.2. *Find all differentiable solutions $f : \mathbb{R} \to \mathbb{R}$ of the exponential Cauchy equation (5.8).*

Solution. From (5.8) we find that $f(0) = 1$. Differentiating (5.8) with respect to x we find

$$f'(x+y) = f'(x)\,f(y)\,.$$

Now we set $x = 0$ and $f'(0) = c$. Then

$$f'(y) = c\,f(y) \Rightarrow \int_1^f \frac{df}{f} = c\int_0^y dy \Rightarrow f(y) = a^y\,,$$

where we set $a = e^c$. □

Problem 12.3. *Find all differentiable solutions $f : \mathbb{R} \to \mathbb{R}$ of the Cauchy III equation (5.14).*

Solution. From (5.14) we find that $f(1) = 0$. Differentiating (5.14) with respect to x we find

$$y\,f'(x\,y) = f'(x)\,.$$

Now we set $x = 1$ and $f'(1) = c$. Then

$$y\,f'(y) = c \Rightarrow \int_0^f df = c\int_1^y \frac{dy}{y} \Rightarrow f(y) = c\ln y = \log_\gamma y\,,$$

where we set $c = 1/\ln\gamma$. □

Problem 12.4. *Find all differentiable solutions $f : \mathbb{R} \to \mathbb{R}$ of the power Cauchy equation (5.20).*

Solution. From (5.20) we find that $f(1) = 1$. Then differentiating (5.20) with respect to x we find

$$y\, f'(x\, y) \ = \ f'(x)\, f(y)\ .$$

Now we set $x = 1$ and $f'(1) = c$. Then

$$y\, f'(y) \ = \ c\, f(y) \ \Rightarrow \ \int_0^f \frac{df}{f} \ = \ c \int_1^y \frac{dy}{y} \ \Rightarrow \ \ln f(y) \ = \ c\, \ln y \ \Rightarrow \ f(y) \ = \ y^c\ . \qquad \square$$

Problem 12.5. *Find all differentiable solutions $f : \mathbb{R} \to \mathbb{R}$ of the Jensen functional equation (5.26).*

Solution. Differentiating (5.26) with respect to x we find

$$\frac{1}{2} f'\left(\frac{x+y}{2}\right) \ = \ \frac{f'(x)}{2}\ .$$

Now we set $x = 0$ and $f'(0) = c$. Then

$$f'(y) \ = \ c \ \Rightarrow \ f(y) \ = \ c\, y + b\ . \qquad \square$$

Problem 12.6. *Find all differentiable solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(x+y) + f(x-y) \ = \ 2\, f(x)\ . \tag{12.1}$$

Solution. Differentiating (12.1) with respect to x we find

$$f'(x+y) + f'(x-y) \ = \ 2\, f'(x)\ .$$

Differentiating (12.1) with respect to y we find

$$f'(x+y) - f'(x-y) \ = \ 0\ .$$

Adding the last two equations

$$f'(x+y) \ = \ f'(x)\ .$$

Now if we set $x = 0$ and $f'(0) = c$ we finally find

$$f'(y) \ = \ c\ ,$$

or

$$f(y) \ = \ c\, y + b\ . \qquad \square$$

Problem 12.7. *Find all differentiable solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(x+y) + f(x-y) \;=\; 2\, f(x)\, f(y)\,. \tag{12.2}$$

Solution. Differentiating (12.2) with respect to x we find

$$f'(x+y) + f'(x-y) \;=\; 2\, f'(x)\, f(y)\,. \tag{12.3}$$

Note that the right hand side is differentiable with respect to y. Then the left hand side is too. Therefore f is twice-differentiable.

Differentiating equation (12.3) with respect to x we find

$$f''(x+y) + f''(x-y) \;=\; 2\, f''(x)\, f(y)\,.$$

Differentiating (12.2) with respect to y twice we find

$$f''(x+y) + f''(x-y) \;=\; 2\, f(x)\, f''(y)\,.$$

The left hand sides of the last two equations are equal. Then

$$f''(x)\, f(y) \;=\; f(x)\, f''(y)\,.$$

Now we set $y = 0$. Then either $f(0) = 0$ or $f(0) \neq 0$. If $f(0) = 0$, we also have $f(x) \equiv 0$. If $f(0) \neq 0$, then set $f''(0)/f(0) = \pm\omega^2$. The function f is a solution of the differential equation

$$f''(x) \pm \omega^2 f(x) \;=\; 0\,,$$

with solutions

$$\begin{aligned} f(x) &= A + B\,x\,, \\ f(x) &= A\, \cos \omega x + B\, \sin \omega x\,, \\ f(x) &= A\, \cosh \omega x + B\, \sinh \omega x\,. \end{aligned}$$

The constants A and B can be determined easily using conditions found from the defining functional equation. Setting $y = 0$ in (12.2), we find $f(x)(f(0) - 1) = 0$. If the function is not identically zero, it must satisfy $f(0) = 1$. Then, differentiating (12.2) with respect to y and setting $x = 0$ we find

$$f'(y) - f'(-y) \;=\; 2f'(y) \;\Rightarrow\; f'(y) + f'(-y) \;=\; 0\,.$$

Then setting $x = 0$ in (12.3), we find $0 = f(y)\, f'(0)$, i.e. $f'(0) = 0$. These conditions $f(0) = 1$ and $f'(0) = 0$ require that $B = 0$ and $A = 1$. Therefore

$$\begin{aligned} f(x) &= 1\,, \\ f(x) &= \cos \omega x\,, \\ f(x) &= \cosh \omega x\,, \end{aligned}$$

are the solutions. □

Problem 12.8 (Putnam 1963)**.** *Find every twice-differentiable real-valued function* $f : \mathbb{R} \to \mathbb{R}$ *satisfying the functional equation*

$$f(x+y)\, f(x-y) \;=\; f(x)^2 - f(y)^2\,, \quad \forall x, y \in \mathbb{R}\,.$$

Solution. The identically vanishing function $f(x) \equiv 0$ is an obvious solution. We shall search for solutions that are not identically vanishing. Therefore, there is an $x_0 \in \mathbb{R}$ such that $f(x_0) \neq 0$.

Differentiating with respect to x the given equation we find

$$f'(x+y)\, f(x-y) + f(x+y)\, f'(x-y) \;=\; 2 f(x)\, f'(x)\,.$$

Then differentiating with respect to y

$$f''(x+y)\, f(x-y) - f(x+y)\, f''(x-y) \;=\; 0\,.$$

We make the change of variables $u = x + y$ and $v = x - y$:

$$f''(u)\, f(v) \;=\; f(u)\, f''(v)\,.$$

We set $v = x_0$. The quantity $f''(x_0)/f(x_0)$ may be positive, negative, or zero. We set it equal to $+b^2, -b^2, 0$ respectively:

$$\begin{aligned} f''(u) - b^2 f(u) &= 0\,, \\ f''(u) + b^2 f(u) &= 0\,, \\ f''(u) &= 0\,. \end{aligned}$$

We can find respectively the solutions

$$\begin{aligned} f(u) &= A\, \sin(bu) + B\, \cos(bu)\,, \\ f(u) &= A\, \sinh(bu) + B\, \cosh(bu)\,, \\ f(u) &= A\, u + B\,. \end{aligned}$$

Setting $x = y = 0$ in the defining relation we find that $f(0) = 0$. This implies that $B = 0$. And the final solutions are

$$\begin{aligned} f(u) &= A\, \sin(bu)\,, \\ f(u) &= A\, \sinh(bu)\,, \\ f(u) &= A\, u\,. \end{aligned}$$

□

Problem 12.9. *Find every twice-differentiable real-valued function $f : \mathbb{R} \to \mathbb{R}$ satisfying the functional equation*

$$f(x+y) + f(x-y) \;=\; 2\, f(x)\, g(y)\,, \quad \forall x, y \in \mathbb{R}\,.$$

Solution. If f is the identically vanishing function, $f(x) \equiv 0$, then $g(x)$ is arbitrary. We shall search for solutions that are not identically vanishing. Therefore, there is an $x_0 \in \mathbb{R}$ such that $f(x_0) \neq 0$ and a $y_0 \in \mathbb{R}$ such that $g(y_0) \neq 0$.

Differentiating twice the given equation with respect to x we find

$$f''(x+y) + f''(x-y) \;=\; 2f''(x)\, g(y)\,.$$

Now, differentiating twice the given equation with respect to y to find

$$f''(x+y) + f''(x-y) \;=\; 2f(x)\, g''(y)\,.$$

Comparing the last two equations

$$f''(x)\, g(y) \;=\; f(x)\, g''(y)\,. \tag{12.4}$$

We now set $x = x_0, y = y_0$:

$$\frac{f''(x_0)}{f(x_0)} \;=\; \frac{g''(y_0)}{g(y_0)}\,.$$

Each ratio may be positive, negative, or zero. We set it equal to $+b^2, -b^2, 0$ respectively.

Setting $x = x_0$ in (12.4), we find

$$\begin{aligned} g''(y) - b^2 g(y) &= 0\,,\\ g''(y) + b^2 g(y) &= 0\,,\\ g''(y) &= 0\,. \end{aligned}$$

We can find respectively the solutions

$$\begin{aligned} g(y) &= A_2 \sin(by) + B_2 \cos(by)\,,\\ g(y) &= A_2 \sinh(by) + B_2 \cosh(by)\,,\\ g(y) &= A_2\, y + B_2\,. \end{aligned}$$

Using the $y = y_0$ in (12.4), we find similarly

$$\begin{aligned} f(x) &= A_1 \sin(bx) + B_1 \cos(bx)\,,\\ f(x) &= A_1 \sinh(bx) + B_1 \cosh(bx)\,,\\ f(x) &= A_1\, x + B_1\,. \end{aligned}$$

Inserting the results in the defining equation, we find the final solutions:

$$\begin{aligned} f(x) &= A_1 \sin(bx) + B_1 \cos(bx)\,, \quad g(x) = A_2 \sin(bx) + B_2 \cos(bx)\,,\\ f(x) &= A_1 \sinh(bx) + B_1 \cosh(bx)\,, \quad g(x) = A_2 \sinh(bx) + B_2 \cosh(bx)\,,\\ f(x) &= A_1\, x + B_1\,, \quad g(x) = 1\,. \end{aligned}$$

□

Comment. The reader has recognized of course that the previously three functional equations we solved are the D'Alembert-Poisson I (7.21), D'Alembert-Poisson II (7.25) and the Wilson I equation (8.25). Here we have solved them by imposing on the solutions stronger conditions than continuity. Note, however, that the set of solutions are the same as those for continuous functions. We will return to this issue in Section 12.3. □

Problem 12.10. *Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function such that*

$$f(x + yf(x)) \;=\; f(x)\, f(y)\,.$$

Solution. Setting $x = y = 0$ we find that either $f(0) = 0$ or $f(0) = 1$. If $f(0) = 0$ then setting $x = 0$ the defining equation implies that $f(x)$ vanishes identically. Therefore, we shall investigate the second case $f(0) = 1$.

Differentiating the defining equation with respect to y we find

$$f'(x + yf(x))\, f(x) \;=\; f(x)\, f'(y)\,.$$

In this equation, let $x = x_0$, where x_0 is one of the points for which f has a non-zero value. Then

$$f'(x_0 + yf(x_0)) \;=\; f'(y)\,. \tag{12.5}$$

Differentiating now the defining equation with respect to x we find

$$f'(x + yf(x))\,(1 + yf'(x)) \;=\; f'(x)\, f(y)\,.$$

In this equation, let $x = x_0$. Then

$$f'(x_0 + yf(x_0))\,(1 + yf'(x_0)) \;=\; f'(x_0)\, f(y)\,.$$

Using the result (12.5), we find

$$f'(y)\,(1 + yf'(x_0)) \;=\; f'(x_0)\, f(y)\,.$$

or

$$f'(y) - \frac{a}{1 + ay}\, f(y) \;=\; 0\,,$$

where we have set $a = f'(x_0)$. The previous equation may equivalently be written as

$$\frac{d}{dy}\left(\frac{f(y)}{1 + ay}\right) \;=\; 0\,.$$

Upon integration we find

$$f(y) \;=\; 1 + a\, y\,.$$

□

12.2 Analytic Functions

Problem 12.11. *Find an analytic solution $f : [0,1) \to \mathbb{R}$ of the functional equation*

$$f(x) + f(x^2) \;=\; x\,, \quad \forall x\,. \tag{12.6}$$

Solution. Since f is analytic, it has a convergent Taylor series for all $x \in [0,1)$. So let

$$f(x) \;=\; \sum_{n=0}^{\infty} a_n\, x^n\,.$$

Then

$$f(x^2) \;=\; \sum_{n=0}^{\infty} a_n\, x^{2n}\,.$$

Substituting in equation (12.6), we find

$$\begin{aligned} x &= \sum_{n=0}^{\infty} a_n\, x^n + \sum_{n=0}^{\infty} a_n\, x^{2n} \\ &= \sum_{k=0}^{\infty} a_{2k}\, x^{2k} + \sum_{k=0}^{\infty} a_{2k+1}\, x^{2k+1} + \sum_{n=0}^{\infty} a_n\, x^{2n}\,, \end{aligned}$$

or

$$\sum_{k=0}^{\infty} (a_{2k} + a_k)\, x^{2k} + (a_1 - 1)\, x + \sum_{k=1}^{\infty} a_{2k+1}\, x^{2k+1} \;=\; 0\,.$$

This equation will be true if

$$\begin{aligned} a_{2k} + a_k &= 0\,, \quad k = 0, 1, 2, \cdots\,, \\ a_{2k+1} &= 0\,, \quad k = 1, 2, \cdots\,, \\ a_1 &= 1\,. \end{aligned}$$

From the second equation we see that all coefficients with subscripts a_{2k+1} vanish. From the first equation we see that if k = odd, $a_{2k} = a_k = 0$. If k = even = $2m$, then $a_{2^2 m} = a_{2k} = -a_k = -a_{2m} = a_m$. Now, a_m will vanish unless $m = 2m'$. Then $a_{2^3 m'} = a_{2m'} = -a_{m'}$. Inductively, $a_{2^\ell m''} = (-1)^\ell a_{m''}$. Therefore, we conclude that $a_{2^\ell} = (-1)^\ell$, $\ell = 0, 1, \ldots$ and all other coefficients vanish, that is

$$f(x) \;=\; \sum_{\ell=0}^{\infty} (-1)^\ell\, x^{2^\ell}\,. \qquad \square$$

The solution found above is the only analytic solution of equation (12.6). There are, however, infinitely many solutions for (12.6). If we modify the domain of f slightly the result can change:

Problem 12.12. *Prove that there is no continuous function $f : [0,1] \to \mathbb{R}$ of the functional equation (12.6).*

Solution. Using the defining equation recursively, we find

$$\begin{aligned} f(x) &= x - f(x^2) \\ &= x - \left(x^2 - f(x^4)\right) \\ &= x - x^2 + \left(x^4 - f(x^8)\right) \\ &\dots \\ &= x - x^2 + x^4 - x^8 + \cdots + (-1)^n\left(x^{2^n} - f\left(x^{2^{n+1}}\right)\right) \end{aligned}$$

This result is true for all $x \in [0,1]$.

Setting $x = 0$ in equation (12.6), we find $f(0) = 0$. By the continuity of f, for any any sequence $a_n \to 0$,

$$\lim_{n\to\infty} f(a_n) = 0 .$$

For any $x \in [0,1)$, we can take

$$a_n = x^{2^{n+1}} .$$

By taking the limit $n \to \infty$ in the equation of $f(x)$, we then conclude that

$$f(x) = x - x^2 + x^4 - x^8 + \cdots + (-1)^n x^{2^n} + \cdots ,$$

for $x \in [0,1)$ — the result of the last problem with a milder assumption.

It remains to examine the behavior of f at $x = 1$. Setting $x = 1$ in equation (12.6), we find $f(1) = 1/2$. Equation (12.6) can equivalently be written

$$f(\sqrt{x}) + f(x) = \sqrt{x} .$$

Using this equation recursively, we find

$$\begin{aligned} f(x) &= \sqrt{x} - f(\sqrt{x}) \\ &= \sqrt{x} - \sqrt[4]{x} - f(\sqrt[4]{x})\Big) \\ &\dots \\ &= \sqrt{x} - \sqrt[4]{x} + \sqrt[8]{x} + \cdots + (-1)^n\left(\sqrt[2^n]{x} - f\left(\sqrt[2^{n+1}]{x}\right)\right) , \end{aligned}$$

or

$$f(x) + (-1)^n f\left(\sqrt[2^{n+1}]{x}\right) = \sqrt{x} - \sqrt[4]{x} + \sqrt[8]{x} + \cdots + (-1)^n \sqrt[2^n]{x} .$$

Setting $x = 1$ in equation (12.6), we find $f(1) = 1/2$. By the continuity of f, for any any sequence $b_n \to 1$,

$$\lim_{n\to\infty} f(b_n) = \frac{1}{2} .$$

For any $x \in (0,1]$, we can take

$$b_n = \sqrt[2^{n+1}]{x} .$$

However, by attempting to take the limit, the term

$$(-1)^n f\left(\sqrt[2^{n+1}]{x}\right)$$

oscillates: the even subsequence b_{2n} gives $f(1/2)$ and the odd subsequence b_{2n+1} gives $-f(1/2)$. Therefore we cannot construct a continuous function with domain [0,1]. □

Problem 12.13 (Putnam 1972)**.** *Show that the power series representation for the series*

$$\sum_{k=0}^{\infty} \frac{x^k (x-1)^{2k}}{k!}$$

cannot have three consecutive zero coefficients.

Solution. We first observe that the given series is an analytic function

$$\sum_{k=0}^{\infty} \frac{(x(x-1)^2)^k}{k!} = e^{x(x-1)^2} .$$

We will set

$$f(x) = e^{g(x)} , \quad g(x) = x^3 - 2x^2 + x .$$

The power series of $f(x)$ sought is its Taylor series:

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k .$$

We must therefore show that, for a random $k = 1, 2, \ldots$, the derivatives $f^{(k-1)}(0)$, $f^{(k)}(0)$, $f^{(k+1)}(0)$ are not simultaneously zero.

We have

$$f'(x) = f(x)\, g'(x) ,$$

and

$$f^{(k)}(x) = \frac{d^{k-1}}{dx^{k-1}} f'(x) = \frac{d^{k-1}}{dx^{k-1}} f(x)\, g'(x) = \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} f^{(\ell)}(x)\, g^{(k-\ell)}(x) .$$

Since $g(x)$ is a polynomial of degree 3, its derivatives from the fourth and above all vanish. Then

$$f^{(k)}(x) \;=\; f^{(k-1)}(x)\,g'(x) + (k-1)\,f^{(k-2)}(x)\,g''(x) + \frac{(k-1)(k-2)}{2}\,f^{(k-3)}(x)\,g'''(x)\,.$$

So finally $f(0) = 1$, $f'(0) = 1$, $f''(0) = -3$ and

$$f^{(k)}(0) \;=\; f^{(k-1)}(0) - 4(k-1)\,f^{(k-2)}(0) + 3(k-1)(k-2)\,f^{(k-3)}(0)\,, \quad k \geq 3\,.$$

We can now prove the statement as follows: If for any k all three derivatives $f^{(k-1)}(0)$, $f^{(k)}(0)$, $f^{(k+1)}(0)$ were zero, then, since the derivative $f^{(k+2)}(0)$ is a linear combination of them, $f^{(k+2)}(0) = 0$ and then $f^{(k+3)}(x)(0) = 0$ and so on to infinity. That is, the Taylor series would be reduced to a finite sum (polynomial) which contradicts the initial expression containing powers x^k, with k larger than any given finite number. We must therefore conclude that the three derivatives $f^{(k-1)}(0)$, $f^{(k)}(0)$, $f^{(k+1)}(0)$ cannot vanish simultaneously for any $k > 0$. □

12.3 Stronger Conditions as a Tool

As has been already mentioned in Section 12.1, even if we are searching for continuous solutions of a functional equation, it is useful to assume that the function is differentiable (one or more times) or even analytic. This allows us to use the vast number of tools from the theory of differential equations. In fact, it allows a uniform treatment of virtually all functional equations, a topic in which a universal approach is missing.

After differential (or analytic) solutions of a functional equation are found, they are examined for uniqueness. This idea is demonstrated by solving again the functional equation (5.29) — Problem 12.14 below. In other cases, one can show that the solutions must necessarily be differentiable (as many times as necessary). This idea has already been presented in Problem 12.1. It is reinforced with Problem 12.15 below.

Problem 12.14. *Find all continuous solutions $f : \mathbb{R} \to \mathbb{R}$ of the functional equation*

$$f(x+y) \;=\; f(x) + f(y) + a\,(1-b^x)\,(1-b^y)\,,$$

where a, b are real constants and $b > 0$.

Solution. Assume that $f_p(x)$ is a doubly differentiable particular solution of (5.29). Then differentiating (5.29) with respect to x and then with respect to y, we find

$$f_p''(x+y) \;=\; a\,(b^x\,\ln b)\,(b^y\,\ln b) \;=\; a\,b^{x+y}\,\ln^2 b\,.$$

Setting $u = x + y$ and integrating twice, we find

$$f_p(x) \;=\; a\,b^x + c_1\,x + c_2\,,$$

where c_1, c_2 some constants. For $x = y = 0$, equation (5.29) gives $f_p(0) = 0$ which, in turn, implies $c_2 = -a$. Therefore

$$f_p(x) \;=\; a\,(b^x - 1) + c_1\,x\,.$$

Let $f(x)$ be any other continuous solution of the functional equation. Then the function $g(x) = f(x) - f_p(x)$ is a continuous function and satisfies

$$g(x+y) \;=\; g(x) + g(y)\,.$$

We thus conclude that $g(x) = k\,x$ with k some real constant and

$$f(x) \;=\; g(x) + f_p(x) \;=\; a\,(b^x - 1) + c\,x\,,$$

with $c = k + c_1$. □

Comment. It is apparent that the current method is more straightforward that the transformation (5.30) since it involves no guessing. In general, it is quite hard to guess the transformation that reduces a complicated equation into a simpler one. □

Problem 12.15. *In Section 8.6, we wrote down the solutions for Wilson's functional equations for continuous functions without calculation. In Section 12.1, we worked out Wilson's first equation (8.25) under the condition of twice-differentiable functions. The solutions are identical. If one was to use differentiation to prove the results of Section 8.6, he should have to establish that no continuous (but not twice-differentiable) solution has been lost in the calculation. Therefore, prove that the original condition of continuity implies the condition of double differentiability.*

Solution. Let

$$F(x) \;=\; \int_0^x f(u)\,du\,, \quad G(x) \;=\; \int_0^x g(u)\,du\,.$$

With this convention, $F(0) = G(0) = 0$. The functions $F(x)$ and $G(x)$ are obviously differentiable with $F'(x) = f(x)$ and $G'(x) = g(x)$.

We now integrate equation (8.25) with respect to y from 0 to y:

$$\int_0^y f(x+z)\,dz + \int_0^y f(x-z)\,dz \;=\; 2f(x)\int_0^y g(z)\,dz\,,$$

or

$$[F(x+y) - F(x)] + [F(x) - F(x-y)] \;=\; 2f(x)\,G(y)\,.$$

This is of course

$$F(x+y) - F(x-y) \;=\; 2f(x)\,G(y)\,.$$

Since the left hand side of this equation is differentiable, the right hand side is differentiable too. This, in turn, implies that $f(x)$ is differentiable. We can thus differentiate the last equation with respect to x to find:

$$f(x+y) - f(x-y) \;=\; 2f'(x)\,G(y)\,.$$

Again, since the left hand side of this equation is differentiable with respect to x, the right hand side is differentiable too which implies that $f''(x)$ exists. Returning to the original equation, since $f(x)$ is twice-differentiable, it is easily seen that $g(x)$ is also twice-differentiable.

Then the solution of (8.25) presented in Section 12.1 gives all continuous solutions. $\square$

Chapter 13

Functional Equations for Polynomials

13.1 Fundamentals

Polynomials hold a special place in mathematics. Functional equations satisfied by polynomials can be solved, of course, by all methods discussed so far and all ones to be discussed in the following chapters. However, knowing that the solution to a functional equation is a polynomial can, as a rule, simplify the solution process since the requirement that the solution be a polynomial is a very restrictive one.

In particular, the following properties are quite useful:

Theorem 13.1. *If $P(x)$ is a polynomial of degree n and $Q(x)$ is a polynomial of degree m, the sum $P(x) + Q(x)$ is a polynomial of degree $\max\{n, m\}$, the product $P(x)Q(x)$ is a polynomial of degree $n + m$, and the composition $P(Q(x))$ is a polynomial of degree nm.*

Proof. Straightforward. □

Theorem 13.2 (Binomial Theorem)**.** *For any x, y and $n \in \mathbb{N}^*$,*

$$(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} .$$

Proof. Easy by induction. It can also be seen as a special case of equation (13.1) below. □

Theorem 13.3 (Fundamental Theorem of Algebra)**.** *A polynomial $P(x)$ of degree $n \geq 1$ has at least one root in $\mathbb{C}$.*

The first proof of this theorem was given by D'Alembert. However, in his proof, D'Alembert had used one assumption without proof: a continuous function, defined on a closed and bounded set, acquires somewhere in the set a minimum. It took another hundred years for this assumption to be resoved.

Let r be the root of $P(x)$, that is $P(r) = 0$. If $P(x)$ is divided by $x - r$ which is of degree 1, then the remainder R will be of degree 0:

$$P(x) \;=\; (x - r)\, Q(x) + R\;,$$

where the quotient $Q(x)$ is of degree $n - 1$. Setting $x = r$ in this equation, we see that $R = 0$. Therefore we have shown the following:

Theorem 13.4. *A polynomial $P(x)$ is divisible by the binomial $x - a$ iff $P(a) = 0$.*

Also, if r is a root of $P(x)$, then

$$P(x) \;=\; (x - r)\, Q(x)\;.$$

Theorem 13.5. *A polynomial $P(x)$ of degree $n \geq 1$ has exactly n roots in $\mathbb{C}$.*

Proof. As proven above, $P(x)$ has at least one root r_1 and it can be factorized in the form

$$P(x) \;=\; (x - r_1)\, P_{n-1}(x)\;,$$

where $P_{n-1}(x)$ is of degree $n - 1$. If $n - 1 \geq 1$, then $P_{n-1}(x)$ has at least one root r_2 and it can be factorized in the form

$$P_{n-1}(x) \;=\; (x - r_2)\, P_{n-2}(x)\;,$$

where $P_{n-2}(x)$ is of degree $n-2$. Continuing this way, we finally reach a quotient polynomial $P_0(x)$ of degree 0, that is one that is a constant:

$$P_1(x) \;=\; (x - r_n)\, P_0(x) \;=\; (x - r_n)\, A\;.$$

Combining all equations we see that $P(x)$ factorizes in the form

$$P(x) \;=\; A\,(x - r_1)(x - r_2)\ldots(x - r_n)\;.$$

Obviously, it has exactly n roots. □

Notice that the n roots may not be distinct. Instead, root r_1 might appear m_1 times, root r_2 might appear m_2 times, and so on. The numbers $m_1, m_2, \ldots$ are known as **multiplicities** of the roots and according to the previous theorem $m_1 + m_2 + \cdots = n$. A root of multiplicity 1 is called **simple**.

Theorem 13.6. *A polynomial $P(x)$ of degree n is identically zero if it vanishes for (at least) $n + 1$ distinct values of x.*

Proof. Let $r_1, r_2, \ldots, r_n$ be n distinct values of x at which $P(x)$ vanishes. According to Theorem 13.5,

$$P(x) \;=\; A(x - r_1)(x - r_2)\ldots(x - r_n)\;.$$

Let r_{n+1} be the $(n + 1)$-th value of x for which $P(x)$ vanishes. Then

$$P(r_{n+1}) \;=\; A(r_{n+1} - r_1)(r_{n+1} - r_2)\ldots(r_{n+1} - r_n)\;.$$

Since r_{n+1} is distinct from $r_1, r_2, \ldots, r_n$, none of the binomial factors $r_{n+1} - r_i$, $i = 1, 2, \ldots, r_n$ vanish. Therefore it must be $A = 0$ which implies that $P(x) \equiv 0$. □

Theorem 13.7. *Two polynomials $P(x)$ and $Q(x)$ of degree n and $m \leq n$ respectively are equal if they obtain equal values for (at least) $n+1$ distinct values of x.*

Proof. We set $R(x) = P(x) - Q(x)$. This polynomial is of degree n and it vanishes for (at least) $n+1$ distinct values of x. According to the previous theorem, it vanishes identically. Therefore $P(x) \equiv Q(x)$. □

Theorem 13.8. *The only periodic polynomial*

$$P(x+T) \;=\; P(x)\,,$$

for some constant T, is the constant polynomial $P(x) = c$.

Proof. Let $P(0) = c$. By the periodicity of $P(x)$, it is true that

$$P(0) \;=\; P(T) \;=\; P(2T) \;=\; \dots \;=\; c\,.$$

Therefore the polynomial $P(x)$ and the constant polynomial $Q(x) = c$ take the same values at an infinite number of points. Therefore, they must be equal to each other. □

13.2 Symmetric Polynomials

For a quadratic polynomial

$$P(x) \;=\; a\,x^2 + b\,x + c\,,$$

the formulæ

$$r_1 + r_2 \;=\; -\frac{b}{a}\,, \quad r_1 r_2 \;=\; \frac{c}{a}\,.$$

providing the sum and product of its two roots are well known. They can be proved by writing

$$P(x) \;=\; a\,(x - r_1)\,(x - r_2)\,,$$

expanding the product, and finally matching the coefficients with the coefficients of the original expression.

These formulæ can be generalized for any polynomial.

Theorem 13.9. *Given the roots $r_1, r_2, \dots, r_n$ of a polynomial*

$$P(x) \;=\; a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0\,,$$

the following formulæ are true

$$\begin{aligned}
\sum_{i=1}^{n} r_i &= -\frac{a_{n-1}}{a_n}\,, \\
\sum_{j>i=1}^{n} r_i r_j &= +\frac{a_{n-2}}{a_n}\,, \\
\sum_{k>j>i=1}^{n} r_i r_j r_k &= -\frac{a_{n-3}}{a_n}\,, \\
&\dots \\
\prod_{i=1}^{n} r_i &= (-1)^n\,\frac{a_0}{a_n}\,,
\end{aligned}$$

and they are known[1] *as* ***Viète's formulæ****.*

Perhaps the proof is obvious to the reader. I give it in the following lines while I am introducing some additional material too. I start by extending the binomial theorem to a product of different factors:

$$\begin{aligned}
x + x_1 &= x + x_1\,, \\
(x + x_1)(x + x_2) &= x^2 + (x_1 + x_2)x + x_1 x_2\,, \\
(x + x_1)(x + x_2)(x + x_3) &= x^3 + (x_1 + x_2 + x_3)x^2 + (x_1x_2 + x_2x_3 + x_3x_1)x + x_1x_2x_3\,, \\
\dots
\end{aligned}$$

and by induction

$$\prod_{i=1}^{n}(x + x_i) = \sum_{i=0}^{n} e_i(x)\,x^{n-i}\,, \tag{13.1}$$

where

$$e_i(x_1, x_2, \dots, x_n) = \sum_{1\le k_1<k_2<\dots<k_i\le n} x_{k_1}x_{k_2}\dots x_{k_i}$$

is known as the i-th **elementary symmetric polynomial** in n variables. (The term 'symmetric' is defined formally below.) If $x_1 = x_2 = \dots = x_n$ we recover the binomial theorem.

Viète's formulæ can be written as

$$e_i(r_1, r_2, \dots, r_n) = (-1)^i\,\frac{a_{n-i}}{a_n}\,,$$

and they are an immediate consequence of equation (13.1) and the factorization of the polynomial $P(x)$:

$$P(x) = a_n\,(x - r_1)(x - r_2)\dots(x - r_n)\,.$$

[1] After the French mathematician François Viète. Often he is mentioned with the Italian version of his name, Franciscus Vieta.

Example 13.1. It is well known that the quadratic polynomial $P(x) = ax^2 + bx + c$ does not have real roots if $b^2 < 4ac$. Imagine that there was no explicit formula for the roots of the quadratic polynomial. Could we have derived the same conclusion in such a case? The answer is yes. From the formulæ

$$r_1 + r_2 = -\frac{b}{a}, \quad r_1 r_2 = \frac{c}{a},$$

we see that

$$(r_1 + r_2)^2 - 2r_1 r_2 = \frac{b^2 - 2ac}{a^2},$$

or

$$r_1^2 + r_2^2 = \frac{b^2 - 2ac}{a^2}.$$

If the roots are real, the left hand side is non-negative. Therefore, if $b^2 < 2ac$ there cannot be real roots. However, this inequality, it is not the most strict. We can compute[2]

$$(r_1 - r_2)^2 = \frac{b^2 - 4ac}{a^2}.$$

From this we can get that there cannot be real roots if $b^2 < 4ac$ that improves the previous result.

You can use this idea in higher order polynomials. Just to get experience, let's do it for the cubic polynomial:

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

Then

$$r_1 + r_2 + r_3 = -\frac{b}{a}, \quad r_1 r_2 + r_2 r_3 + r_3 r_1 = \frac{c}{a},$$

and

$$\begin{aligned} r_1^2 + r_2^2 + r_3^2 &= \frac{b^2 - 2ac}{a}, \\ (r_1 - r_2)^2 + (r_2 - r_3)^2 + (r_3 - r_1)^2 &= \frac{2b^2 - 6ac}{a^2}. \end{aligned}$$

If $b^2 < 3ac$, a cubic polynomial cannot have three real roots.

Now the solution to the following problem should be obvious to you.

Problem 13.1 (USA 1983)**.** *Prove that the roots of*

$$x^5 + bx^4 + cx^3 + dx^2 + ex + f = 0$$

cannot all be real if $2b^2 < 5c$.

[2]We see that for a quadratic polynomial, its discriminant is given in terms of the difference of its roots: $D = a^2(r_1 - r_2)^2$. In general, for a polynomial of degree n, we define its discriminant by $D = a_n^{2(n-1)} \prod_{i=1}^{n} \prod_{j=i+1}^{n} (r_i - r_j)^2$.

In the previous example we used the expression $P(x, y, z) = x^2 + y^2 + z^2$ which is a polynomial itself with the additional property that it remains unchanged under exchanges of its variables. Such a polynomial is called *symmetric*:

Definition 13.1. A **symmetric polynomial** $S(x_1, x_2, \ldots, x_n)$ in n variables is a polynomial of the n variables that satisfies

$$S(x_1, x_2, \ldots, x_n) \;=\; S(x_{\sigma_1}, x_{\sigma_2}, \ldots, x_{\sigma_n}) \,,$$

for any permutation $\sigma_1, \sigma_2, \ldots, \sigma_n$ of $1,2,\ldots,n$.

Theorem 13.10. *Any symmetric polynomial can be expressed as a polynomial in the elementary symmetric polynomials.*

From what has been discussed above, it is immediate that

Corollary. *Any symmetric polynomial $S(r_1, r_2, \ldots, r_n)$ of the roots of a polynomial $P(x)$ can be expressed in terms of the coefficients of the polynomial $P(x)$.*

The previous corollary explains why, in Example 13.1, I used symmetric polynomials

$$\begin{aligned} P(x, y, z) &= x^2 + y^2 + z^2 \,, \\ P(x, y, z) &= (x - y)^2 + (y - z)^2 + (z - x)^2 \,, \end{aligned}$$

to draw conclusions about the roots and not polynomials of the form

$$P(x, y, z) \;=\; (x - y)^2 + y^2 + z^2$$

for example. The former are written easily in terms of the coefficients of the polynomial through Viète's formulæ.

13.3 More on the Roots of Polynomials

Most of the theorems presented above on polynomials dealt with their roots. So, a section entitled "more on the roots of polynomials" might be an overkill. But there are some additional comments and results regarding the roots of a polynomial that are worth mentioning.

Theorem 13.11. *(a) If r is a simple root of the polynomial $P(x)$, then it is not a root of its derivative $P'(x)$.*

(b) If r is a root of multiplicity m, $m \geq 2$, then it is a root of $P'(x)$ with multiplicity $m - 1$.

Proof. Let

$$P(x) \;=\; (x - r)^m \, Q(x) \,,$$

with $m \geq 1$. Since m gives the maximal exponent for the factor $(x - r)^m$ which divides $P(x)$, $x - r$ does not divide $Q(x)$ and therefore $Q(r) \neq 0$.

If we differentiate the above equation, we find:

$$P'(x) \;=\; (x-r)^{m-1}\;[m\,Q(x) + (x-r)\,Q'(x)]\;.$$

Obvioulsy, r is a root of $P'(x)$ with multiplicity at least $m-1$. Now, $x-r$ does not divide the quotient $q(x) = m\,Q(x) + (x-r)\,Q'(x)$ since $q(r) = m\,Q(r) \neq 0$. Therefore, r is a root of $P(x)$ with multiplicity $m-1$. In particular, if $m = 1$, then it has multiplicity 0 for $P'(x)$ — that is, it is not a root. □

Rolle's theorem provides a useful tool that can be used to draw conclusions about the existence of roots for a polynomial.

Theorem 13.12. *If a and b are two distinct roots of the polynomial $P(x)$, then $P'(x)$ has a simple root in the interval (a, b).*

Proof. Since $P(x)$ is continuous and differentiable in all $\mathbb{R}$, it satisfies the conditions of Rolle's theorem in the interval $[a, b]$. Therefore, there exists an ξ such that $P'(\xi) = 0$, that is a root of $P'(x)$.

To prove that such a root is simple, we will do simple counting. Imagine that $P(x)$ has n real roots $r_1, \ldots, r_k$ with multiplicities $m_1, \ldots m_k$ (the multiplicities may be 1). Then $m_1 + \cdots + m_k = n$. Its derivative $P'(x)$ is a polynomial of degree $n-1$ and has, at least, the following roots: $r_1, \ldots, r_k$ with multiplicities $m_1, \ldots m_k$ (some multiplicities may be 0) and $\xi_1, \xi_2, \ldots, \xi_{k-1}$ in the intervals (r_1, r_2), (r_2, r_3), …, (r_{k-1}, r_k). Therefore, $P'(x)$ has at least $(m_1 - 1) + \cdots + (m_k - 1) + k - 1 = n - k + k - 1 = n - 1$ roots. But this is exactly the numbers of roots $P'(x)$ must have. We thus conclude that all ξ_i, $i = 1, 2, \ldots, k-1$ are simple roots. □

Problem 13.2 (BWMC 1991)**.** *Let a, b, c, d, e be distinct real numbers. Prove that the equation*

$$\begin{aligned} & (x-a)(x-b)(x-c)(x-d) \\ + \;\; & (x-a)(x-b)(x-c)(x-e) \\ + \;\; & (x-a)(x-b)(x-d)(x-e) \\ + \;\; & (x-a)(x-c)(x-d)(x-e) \\ + \;\; & (x-b)(x-c)(x-d)(x-e) \;=\; 0 \end{aligned}$$

has four distinct real solutions.

Solution. The left hand side of the given equation is the derivative $P'(x)$ of the polynomial

$$P(x) \;=\; (x-a)(x-b)(x-c)(x-d)(x-e)$$

which has five distinct real roots a, b, c, d, e. Without loss of generality we will assume that $a < b < c < d < e$. As proved above, $P'(x)$ has a simple root in each of the intervals (a, b), (b, c), (c, d), (d, e) — that is, $P'(x)$ four distinct roots. □

We can use this idea in 'reverse' as the solution to the following problem demonstrates.

Problem 13.3 (USA 1983). *Prove that the roots of*

$$x^5 + bx^4 + cx^3 + dx^2 + ex + f = 0$$

cannot all be real if $2b^2 < 5c$.

Solution. We set $P(x) = x^5 + bx^4 + cx^3 + dx^2 + ex + f$. Now let's assume that $P(x)$ has five real roots. According to the argument of the previous problem, the polynomial

$$P_1(x) = P'(x) = 5x^4 + 4bx^3 + 3cx^2 + 2dx + e$$

will have four real roots.

However, we can repeat the argument for the polynomial $P_1(x)$. Since it has four real roots, the polynomial

$$P_2(x) = P_1'(x) = 20x^3 + 12bx^2 + 6cx + 2d$$

will have three real roots.

And once more: Since $P_2(x)$ has three real roots, the polynomial

$$P_3(x) = P_2'(x) = 60x^2 + 24bx + 6c = 6\,(10x^2 + 4bx + c)$$

will have two real roots. This is a quadratic polynomial and we can decide if it has real roots by computing its discriminant. For the polynomial inside the parenthesis,

$$D = 16b^2 - 40c = 8(2b^2 - 5c)\,.$$

Therefore if $2b^2 < 5c$, $D < 0$ and the polynomial does not have real roots. Therefore, the initial assumption that $P(x)$ has five real roots is not correct. □

Theorem 13.13. *The roots of the polynomial*

$$P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

cannot all be real if

$$(n-1)\,a_{n-1}^2 < 2n\,a_n\,a_{n-2}$$

Proof. If $P(x)$ had n real roots then

$$\begin{aligned}\frac{d^{n-2}}{dx^{n-2}}P(x) &= \frac{n!\,a_n}{2}\,x^2 + (n-1)!\,a_{n-1}\,x + a_{n-2}\,(n-2)! \\ &= (n-2)!\left[\frac{(n-1)n\,a_n}{2}\,x^2 + (n-1)\,a_{n-1}\,x + a_{n-2}\right]\end{aligned}$$

would have two real roots. The discriminant of the quadratic polynomial is

$$D = (n-1)\left[(n-1)a_{n-1}^2 - 2na_na_{n-2}\right]\,.$$

If $D < 0$, equivalently $(n-1)a_{n-1}^2 < 2na_na_{n-2}$, the quadratic polynomial cannot have real roots and therefore $P(x)$ cannot have all real roots. □

The following two lemmas will be useful in what follows.

Lemma 13.1. *A polynomial $P(x)$ has the same number of positive roots or one more than its derivative $P'(x)$.*

Proof. Let $p_1, p_2, \ldots p_\ell$ be distinct positive roots of $P(x)$ with multiplicities $m_1, m_2, \ldots, m_\ell$ respectively. Then $p_1, p_2, \ldots p_\ell$ are positive roots of $P'(x)$ with multiplicities $m_1 - 1, m_2 - 1, \ldots, m_\ell - 1$ respectively and in each of the intervals (p_1, p_2), (p_2, p_3), $\ldots (p_{\ell-1}, p_\ell)$, by Rolle's theorem, there is a root of $P'(x)$. This gives

$$(m_1 - 1) + (m_2 - 1) + \cdots + (m_\ell - 1) + \ell - 1 \;=\; m_1 + m_2 + \cdots + m_\ell - 1$$

positive roots for $P'(x)$.

Then we have the possibilities: (a) $P(x)$ has no other roots. (b) $x = 0$ is a root of $P(x)$. (c) $P(x)$ has a largest negative root r. They imply respectively: (a) $P'(x)$ has no additional positive roots. (b) There is one more root of $P'(x)$ in $(0, p_1)$. (c) $P'(x)$ will have a root in (r, p_1) which may or may not be positive. This concludes the proof. □

Lemma 13.2. *If a_k is the first non-vanishing coefficient of polynomial of degree n with ℓ positive roots, then*

$$sgn(a_k) \;=\; (-1)^\ell \, sgn(a_n) \,.$$

Proof. For the polynomial

$$P(x) \;=\; a_n x^n + \cdots + a_{k+1} x^{k+1} + a_k x^k \;=\; x^k \left(a_n x^{n-k} + \cdots + a_{k+1} x + a_k\right),$$

let $p_1, \ldots p_\ell$ be its positive roots and $r_{\ell+1}, \ldots, r_{n-k}$ the negative ones. Then

$$P(x) \;=\; a_n \, x^k \, (x - p_1) \cdots (x - p_\ell)(x - r_{\ell+1}) \cdots (x - r_{n-k}) \,.$$

From this equation we find

$$a_k \;=\; a_n (-1)^\ell p_1 \cdots p_\ell \, (-r_{\ell+1}) \cdots (-r_{n-k}) \,,$$

from which the advertised relation follows. □

Theorem 13.14 (Descartes's Law of Signs)**.** *If all the coefficients and roots of the polynomial $P(x)$ are real, then the number of positive roots (multiplicities taken into account) is equal to the number of sign changes in the sequence of coefficients of $P(x)$.*

Proof. Without loss of generality, we assume that the coefficient a_n in

$$P(x) \;=\; a_n x^n + \cdots + a_0$$

is positive. Then we notice that if all coefficients up to and including a_{k-1} vanish,

$$P(x) \;=\; a_n x^n + \cdots + a_k x^k \,,$$

we can write the polynomial in the form

$$P(x) = x^k\,(a_n x^{n-k} + \cdots + a_{k+1}x + a_k)\,.$$

The factor x^k contributes neither positive nor negative roots. Therefore $P(x)$ has the same number of positive roots as

$$Q(x) = a_n x^{n-k} + \cdots + a_k$$

and the same sign changes in the sequence of coefficients. However, $Q(x)$ has a non-vanishing constant term. Therefore, without loss of generality, we can assume a non-vanishing constant term $a_0 \neq 0$ for the polynomial $P(x)$.

Given $a_0 \neq 0$ and $a_n > 0$, we will prove the theorem by induction.

•For a polynomial of degree 1, $P(x) = a_1 x + a_0$, $a_0 \neq 0$, $a_1 > 0$, there is a real root $-a_0/a_1$ which is positive if a_0 is negative. Therefore there is one sign change in the coefficients of $P(x)$ and the theorem is true.

•We assume that the assertion is true for any polynomial of degree $n-1$ with a non-vanishing constant term and a positive coefficient for the $(n-1)$-th term.

•We will now show that the assertion is true for a polynomial of degree n with a non-vanishing constant term and a positive coefficient for the n-th term. We consider the derivative of $P(x)$:

$$P'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1\,.$$

Again, without loss of generality, we can assume that $a_1 \neq 0$. $P'(x)$ is of degree $n-1$ and the theorem applies. That is, the number of sign changes in its coefficients equals the number ℓ of its positive roots.

We have seen that $P(x)$ has either the same number of positive roots or one more.

Case (P and P' have the same number of positive roots). In this case, the first non-vanishing coefficient of each polynomial has sign $(-1)^\ell$, where ℓ is the number of positive roots. For $P(x)$ this coefficient is a_0 and for $P'(x)$ is a_1. Looking at the sequences of coefficients for $P(x)$ and $P'(x)$

$$\begin{array}{llll} a_n, & a_{n-1}, \ldots, & a_2,\ a_1,\ a_0, \\ na_n, & (n-1)a_{n-1},\ \ldots, & 2a_2,\ a_1, \end{array}$$

we see that the number of sign changes in the coefficients of $P(x)$ is the same with the number of sign changes in the coefficients of $P'(x)$. Therefore the theorem is valid for a polynomial of degree n is this case.

Case (P' one positive root fewer than P). In this case, the first non-vanishing coefficient of $P'(x)$, a_1, has sign $(-1)^\ell$, and a_0 has sign $(-1)^{\ell+1}$. Looking at the sequences of coefficients for $P(x)$ and $P'(x)$

$$\begin{array}{llll} a_n, & a_{n-1}, \ldots, & a_2,\ a_1,\ a_0, \\ na_n, & (n-1)a_{n-1},\ \ldots, & 2a_2,\ a_1, \end{array}$$

we see that in the first sequence there is an additional change of sign between a_1 and a_0. Therefore, $P(x)$ has $\ell+1$ positive roots and the same number of sign changes in the sequence of its coefficients. This concludes the proof. □

Comment. If $P(x)$ has also complex roots, since they will appear in conjugate pairs, the number of positive roots will be equal to the number of sign changes minus an even number. However, this statement is not automatic. A proof is required which you can provide easily. Finally, notice that if we do not know anything about the roots of the polynomial, then Descartes's law of signs should be stated as follows: *If the coefficients of a polynomial $P(x)$ are real, then the number of its positive roots is equal to the number of sign changes in the sequence of coefficients of $P(x)$ or less than this number by an even integer.*

Problem 13.4 ([59]). *Assuming that the continued radical*

$$\sqrt{1+\sqrt{7+\sqrt{1+\sqrt{7+\cdots}}}}$$

converges, find its value.

Solution. Let

$$\ell = \sqrt{1+\sqrt{7+\sqrt{1+\sqrt{7+\cdots}}}}\,.$$

Then we notice that

$$\ell = \sqrt{1+\sqrt{7+\ell}}\,,$$

or

$$\ell^4 - 2\ell^2 - \ell - 6 = 0\,.$$

The polynomial $P(x) = x^4 - 2x^2 - x - 6$ has real coefficients and the number of sign changes in the sequence of its coefficients is 1. Therefore, $P(x)$ has only positive root. It is easily seen that $P(2) = 0$ and thus $\ell = 2$. □

Theorem 13.15 (Budan). *If all the coefficients and roots of the polynomial $P(x)$ are real, then the number of positive roots (multiplicities taken into account) in the interval (a, b) is equal to the number of sign changes in the sequence of coefficients of $P(x + a)$ minus the number of sign changes in the sequence of coefficients of $P(x + b)$.*

Proof. For the polynomial $Q(x) = P(x + a)$ the number of positive roots N_a will be given by the number of sign changes in the sequence of coefficients of $Q(x)$. But these are exactly the roots of $P(x)$ which are greater than a. Similarly number of roots N_b of $P(x)$ which are greater than b will be given by the number of sign changes in the sequence of coefficients of $P(x + b)$. The difference $N_a - N_b$ is exactly the number of positive roots of $P(x)$ in the interval (a, b). □

Comment. If $P(x)$ has also complex roots, the number of positive roots in the interval (a, b) will be equal to the above difference minus an even number.

Notice that so far, despite all stated theorems, we have not presented any that answers the questions:

1. Given a polynomial $P(x)$ does it have a real root?
2. How many real roots does it have?
3. How many real roots in the interval (a, b)?

These questions are answered by our last theorem due to Charles Sturm. Towards that goal, we must introduce [3] some simple notation and state some auxiliary results.

First, we deal with the possibility of multiple roots. We will show that this case can be reduced to the case of simple roots. Let

$$P(x) = (x - s_1)(x - s_2) \cdots (x - s_\ell)(x - r_1)^{m_1} (x - r_2)^{m_2} \cdots (x - r_k)^{m_k} ,$$

where $s_1, s_2, \ldots, s_\ell$ are simple roots of $P(x)$ and $r_1, r_2, \ldots, r_k$ roots of multiplicities $m_1, m_2, \ldots, m_k$ respectively greater than 1. As we have seen, the derivative of $P'(x)$ will be of the form:

$$P'(x) = (x - a_1)(x - a_2) \cdots (x - a_{\ell-1})(x - r_1)^{m_1 - 1} (x - r_2)^{m_2 - 1} \cdots (x - r_k)^{m_k - 1} ,$$

with $a_1, \ldots, a_{k+\ell}$ simple roots different from $s_1, \ldots, s_\ell$. Therefore,

$$G(x) = (x - r_1)^{m_1 - 1} (x - r_2)^{m_2 - 1} \cdots (x - r_k)^{m_k - 1}$$

is the greatest common divisor of $P(x)$ and $P'(x)$. If we factorize $P(x)$ in the form

$$P(x) = G(x)\, p(x) ,$$

then the equation $P(x) = 0$ is equivalent to

$$G(x) = 0 \quad \text{or} \quad p(x) = 0 .$$

The polynomial $p(x) = 0$ has only simple roots. $G(x)$ may have multiple roots. But repeating the procedure as many times as necessary, we can reduce the initial equation to a sequence of equations involving only simple roots.

For Sturm's theorem, we thus assume that $P(x)$ has only simple roots. Then the greatest common divisor of $P(x)$ and $P'(x)$, which will indicate by $P_1(x)$, is a constant number

[3] Material for Sturm's theorem is based on Dörrie's book [17] which lists this theorem among the top 100 mathematical gems that can be understood with elementary mathematics. The fundemental theorem of algebra and Abel's result on the impossibility to solve algebraically fifth degree equations and higher make also the list. Although I do not fully agree with all items in the list, these three results are certainly rightfully there and [17] is a great book to own and read.

$c \neq 0$. Using Euclid's algorithm for the polynomials $P(x)$ and $P_1(x)$, we will denote the quotients in the successive divisions by $q_i(x)$ and the corresponding remainders by $-P_{i+1}(x)$:

$$\begin{aligned}
P(x) &= P_1(x)\,q_1(x) - P_2(x)\,, \quad \deg P_2 < \deg P_1 = n-1\,,\\
P_1(x) &= P_2(x)\,q_2(x) - P_3(x)\,, \quad \deg P_3 < \deg P_2\,,\\
P_2(x) &= P_3(x)\,q_3(x) - P_4(x)\,, \quad \deg P_4 < \deg P_3\,,\\
\dots\\
P_{s-1}(x) &= P_s(x)\,q_s(x)\,, \quad \deg P_s = 0\,.
\end{aligned}$$

The algorithm terminates at some step, say the s-th step, and $P_s(x) = c$ is the greatest common divisor of $P(x)$ and $P'(x)$. In this way we construct a sequence of polynomials with decreasing degrees,

$$P(x),\ P_1(x),\ P_2(x),\ P_3(x),\ \dots,\ P_s(x).$$

This sequence is called **Sturm chain** for the polynomial $P(x)$. The following three lemmas discuss the properties of Sturm chain.

Lemma 13.3. *Two successive polynomials do not vanish simultaneously at any point.*

Proof. If for some x, $P_i(x) = P_{i+1}(x) = 0$ then from the i-th equation of Euclid's algorithm we see that $P_{i+2}(x) = 0$ which then implies $P_{i+3}(x) = 0$ and so on until the last equation, $P_s(x) = c = 0$. But this is impossible. □

Lemma 13.4. *If r is a root of $P_i(x)$, then $P_{i-1}(r)$ and $P_{i+1}(r)$ have opposite signs.*

Proof. Using the $(i-1)$-th equation, $P_{i-1}(x) = P_i(x)\,q_i(x) - P_{i+1}(x)$, we see immediately that, at $x = r$, $P_{i-1}(r) = -P_{i+1}(r)$. □

Lemma 13.5. *If r is a root of $P(x)$, then there is a neighborhood of r such that $P_1(x)$ maintains the same sign for all points in this neighborhood.*

Proof. If r is a root of $P(x)$, then $P_1(r) \neq 0$. Since $P_1(x)$ is continuous, there is a neighborhood in which it is positive or negative. □

Now, we select a point x such that it is not a root for any of the polynomials in Sturm chain. Then we create the sequence of signs

$$\operatorname{sgn}(P(x)),\ \operatorname{sgn}(P_1(x)),\ \operatorname{sgn}(P_2(x)),\ \operatorname{sgn}(P_3(x)),\ \dots,\ \operatorname{sgn}(P_s(x)),$$

which is known as a **Sturm sign chain**, and let $\sigma(x)$ stand for the number of sign changes[4] in a Sturm sign chain.

[4]Notice that sign changes are the same in the original Sturm chain or in a chain that is formed by the original chain but with the polynomials multiplied by arbitrary positive numbers. As a result, we can always avoid fractional coefficients in the calculations.

Theorem 13.16 (Sturm). *The number of real roots of $P(x)$ in the interval (a, b) is equal to the difference $\sigma(a) - \sigma(b)$.*

Proof. Changes in the value of the function $\sigma(x)$ can happen when one of the polynomials of the chain changes sign. But this means, that x passes through a zero of one of the polynomials. Say that x_0 is a zero of $P_i(x)$. Then there is neighborhood such that, in this neighborhood,

1. $P_i(x)$ does not vanish except at the point x_0 and
2. the functions $P_{i-1}(x)$ and $P_{i+1}(x)$ maintain their sign if $i \neq 0$, or just the function $P_1(x)$ maintains its sign if $i = 0$.

Let's consider the two points $a = x_0 - \epsilon$ and $b = x_0 + \epsilon$, with $\epsilon > 0$, belonging to this neighborhood and look at the signs of the polynomials under discussion.

Case ($i \neq 0$). P_i changes sign as it goes through the point x_0. Therefore $P_i(a)$ and $P_i(b)$ have opposite signs. At the same time $P_{i-1}(a)$ and $P_{i+1}(b)$ have the same sign but exactly opposite of the sign of $P_{i+1}(a)$ and $P_{i-1}(b)$. Consequently, each of the sign chains

$$\operatorname{sgn}(P_{i-1}(a)),\ \operatorname{sgn}(P_i(a)),\ \operatorname{sgn}(P_{i+1}(a)),$$
$$\operatorname{sgn}(P_{i-1}(b)),\ \operatorname{sgn}(P_i(b)),\ \operatorname{sgn}(P_{i+1}(b))$$

has only one sign change and the net contribution to $\sigma(a) - \sigma(b)$ is zero.

Case ($i = 0$). Again, P_0 changes sign as it goes through the point x_0. Therefore $P_0(a)$ and $P_0(b)$ have opposite signs. At the same time $P_1(a)$ and $P_1(b)$ have the same sign. Consequently, among the two pairs

$$\operatorname{sgn}(P_0(a)),\ \operatorname{sgn}(P_1(a)),$$
$$\operatorname{sgn}(P_0(b)),\ \operatorname{sgn}(P_1(b)),$$

one has identical signs and one has opposite signs. In fact, it is always the first pair that has the opposite signs: When $P_1 > 0$, P must be strictly increasing and thus $P(a) < 0$; when $P_1 > 0$, P must be strictly decreasing and thus $P(a) > 0$. The net contribution to $\sigma(a) - \sigma(b)$ is one.

Therefore, when $\sigma(x)$ changes values only when x goes through zeroes of $P(x)$. In particular, it changes by one every time it passes through one such zero. From this, the statement of the theorem follows. □

Example 13.2. For the polynomial $P(x) = x^5 - 3x - 1$ the corresponding Sturm chain is

$$\begin{aligned} P(x) &= x^5 - 3x - 1\,, \\ P_1(x) &= 5x^4 - 3\,, \\ P_2(x) &= 12x + 5\,, \\ P_3(x) &= 1\,. \end{aligned}$$

We will consider the following Sturm sign chains.

1. At $x = -M, M \to +\infty$:
$$-1,\ +1,\ -1,\ +1.$$
2. At $x = -2$:
$$-1,\ +1,\ -1,\ +1.$$
3. At $x = -1$:
$$+1,\ +1,\ -1,\ +1.$$
4. At $x = 0$:
$$-1,\ -1,\ +1,\ +1.$$
5. At $x = 1$:
$$-1,\ +1,\ +1,\ +1.$$
6. At $x = 2$:
$$+1,\ +1,\ +1,\ +1.$$
7. At $x = M, M \to +\infty$:
$$+1,\ +1,\ +1,\ +1.$$

The polynomial $P(x)$ has only three real roots since $\sigma(-\infty) - \sigma(+\infty) = 3$. In fact there is one root in each of the intervals $(-2,-1)$, $(-1,0)$, $(1,2)$ since $\sigma(a) - \sigma(b) = 1$ for any of them. □

13.4 Solved Problems

Polynomials, by evaluating them or by the use of Viète's formulæ can be useful to find the values of finite sums. This idea is demonstrated in the problem that follows.

Problem 13.5. *Prove that*
$$\prod_{k=1}^{n-1} \sin\frac{k\pi}{n} = \frac{n}{2^{n-1}} ,$$
for any $n \geq 2$.

Solution. We start with the polynomial
$$P(x) = x^n - 1 ,$$
having roots
$$\omega_k = e^{i\,\frac{2k\pi}{n}} ,\ \ k = 0, 1, \ldots, n-1 .$$
Since
$$P(x) = (x-1)\,(x^{n-1} + x^{n-2} + \cdots + x + 1) ,$$

the polynomial

$$Q(x) \;=\; x^{n-1} + x^{n-2} + \cdots + x + 1$$

has roots ω_k, $k = 1, \ldots, n-1$. That is

$$Q(x) \;=\; (x - \omega_1)(x - \omega_2) \ldots (x - \omega_{n-1}) \,.$$

Then we notice that

$$\begin{aligned} 1 - \omega_k &= 1 - \cos\frac{2\pi k}{n} + i \sin\frac{2\pi k}{n} \\ &= 2\sin^2\frac{\pi k}{n} + i2\sin\frac{\pi k}{n}\,\cos\frac{\pi k}{n} \\ &= 2ie^{-ik\frac{\pi}{n}}\,\sin\frac{\pi k}{n}\,. \end{aligned}$$

Setting $x = 1$ in $Q(x)$, we find

$$\prod_{i=k}^{n}(1 - \omega_k) \;=\; n\,,$$

which is exactly the identity sought. (Since the roots appear in conjugate pairs, only the 2's and the sine parts of the roots survive. Alternatively, take the magnitude of the two sides.) □

Incidentally, from the previous problem, since $P(\omega_k) = 0$, we also see that, if ω is any of the n-th roots of 1 different from $\omega_0 = 1$,

$$1 + \omega + \omega^2 + \cdots + \omega^{n-1} \;=\; 0\,,$$

another well known identity.

Problem 13.6. *Find all polynomials $P(x)$ such that*

$$x\,P(x-1) \;=\; (x-3)\,P(x)\,, \quad \forall x.$$

Solution. Substituting $x = 0, 1, 2$, successively in the given equation we find that these values are roots of $P(x)$:

$$\begin{aligned} 0\,P(-1) &= -3\,P(0) \;\Rightarrow\; P(0) \;=\; 0\,, \\ 1\,P(0) &= -2\,P(1) \;\Rightarrow\; P(1) \;=\; 0\,, \\ 2\,P(1) &= -1\,P(2) \;\Rightarrow\; P(2) \;=\; 0\,. \end{aligned}$$

Let's set

$$P(x) \;=\; x(x-1)(x-2)\,Q(x)\,,$$

where $Q(x)$ is some polynomial. Substituting in the given equation

$$x(x-1)(x-2)(x-3)Q(x-1) \;=\; x(x-1)(x-2)(x-3)Q(x)\,,$$

or

$$Q(x-1) \;=\; Q(x)\,, \quad \forall x\,.$$

This equation implies that $Q(x)$ is a constant, say c, and thus

$$P(x) \;=\; c\,x(x-1)(x-2)\,. \qquad \square$$

Problem 13.7 (USA 1975)**.** *If $P(x)$ denotes a polynomial of degree n such that*

$$P(k) \;=\; \frac{k}{k+1}\,, \quad k=0,1,\ldots,n\,, \tag{13.2}$$

determine $P(n+1)$.

Solution. The given conditions may be written as

$$(k+1)\,P(k)-k \;=\; 0\,, \quad k=0,1,\ldots,n\,.$$

This implies that the polynomial

$$Q(x) \;=\; (x+1)\,P(x)-x$$

of degree $n+1$ has $0,1,\ldots,n$ as its $n+1$ roots. Therefore:

$$Q(x) \;=\; a\,x(x-1)(x-2)\ldots(x-n)\,.$$

To find a, notice that the definition of $Q(x)$ through $P(x)$ implies that $Q(-1)=1$, while the previous equation gives

$$Q(-1) \;=\; a\,(-1)^{n+1}(n+1)!\,.$$

We thus conclude that

$$a \;=\; \frac{(-1)^{n+1}}{(n+1)!}$$

and

$$P(x) \;=\; \left(\frac{(-1)^{n+1}}{(n+1)!}\,(x-1)(x-2)\ldots(x-n)+1\right)\frac{x}{x+1}\,.$$

Then

$$P(n+1) \;=\; \left(\frac{(-1)^{n+1}}{n+1}+1\right)\frac{n+1}{n+2}\,,$$

or

$$P(n+1) \;=\; \begin{cases} \frac{n}{n+2}\,, & \text{if } n\text{=even}\,,\\ 1\,, & \text{if } n\text{=odd}\,. \end{cases} \qquad \square$$

Problem 13.8. *Find all polynomials $P(x)$ such that*

$$P(F(x)) = F(P(x)), \quad P(0) = 0,$$

where F is some function defined on $\mathbb{R}$ and that satisfies

$$F(x) > x, \forall x \geq 0.$$

Solution. For $x = 0$, the given functional equation gives

$$P(F(0)) = F(0).$$

We set $a_0 = F(0)$ and we thus rewrite this equation as

$$P(a_0) = a_0.$$

Starting with a_0, we now define the sequence

$$a_{n+1} = F(a_n),$$

Since $F(a_n) > a_n$, this sequence is a strictly increasing one. In the defining functional equation, we set $x = a_0$. Then

$$P(F(a_0)) = F(P(a_0)) \Rightarrow P(a_1) = F(a_0) \Rightarrow P(a_1) = a_1.$$

We thus suspect that

$$P(a_n) = a_n, \quad \forall n.$$

Suppose it is true for $n = k$: $P(a_k) = a_k$. Then we set $x = a_k$ in the defining functional equation to find: $P(F(a_k)) = F(P(a_k)) \Rightarrow P(a_{k+1}) = F(a_k) \Rightarrow P(a_{k+1}) = a_{k+1}$. That is, it is also true for $k+1$ and we have proved our claim.

We have thus shown that the polynomial $Q(x) = P(x) - x$ has an infinite number of distinct roots: a_n, with $n = 0, 1, 2, \cdots$. Therefore, it must be the zero polynomial. This implies that

$$P(x) = x. \qquad \square$$

Problem 13.9. *Find all polynomials $P(x)$ for which there exists a polynomial of two variables $Q(u, v)$ such that*

$$P(xy) = Q(x, P(y)).$$

Solution 1. Any polynomial of degree zero $P(x) = a_0$ satisfies the conditions of the problem. We will thus seek polynomials of degree $n \geq 1$.

Let $Q(u, v)$ be a polynomial of degree n_1 with respect to u and degree n_2 with respect to v. Then $Q(x, P(y))$ is of degree n_1 with respect to x and of degree nn_2 with respect to y. Therefore the given equation requires that $n = n_1$ and $n = nn_2$ or $n_1 = n$ and $n_2 = 1$. We can write the polynomial $Q(u, v)$ in the form

$$Q(u, v) \;=\; q_1(u) + q_2(u)\, v \;,$$

where the polynomials $q_1(u), q_2(u)$ are of degree n. The given functional equation then takes the form of a Vincze II functional equation (8.19):

$$P(xy) \;=\; q_1(x) + q_2(x)\, P(y) \;,$$

with solution

$$P(x) \;=\; ax^n + b \;, \quad Q(u, v) \;=\; b\,(1 - u^n) + u^n\, v \;.$$

$\square$

If one does not want to use the result of the Vincze II equation, he can proceed as follows:

Solution 2. Let

$$\begin{aligned} P(x) &= \sum_{k=0}^{n} a_k x^k \;, \\ q_1(u) &= \sum_{k=0}^{n} b_k u^k \;, \\ q_2(u) &= \sum_{k=0}^{n} c_k u^k \;. \end{aligned}$$

Then the given equation translates to

$$\sum_{k=0}^{n} a_k x^k y^k \;=\; \sum_{k=0}^{n} b_k x^k + \sum_{k=0}^{n} \sum_{\ell=0}^{n} c_k a_\ell x^k y^\ell \;,$$

or

$$a_k y^k \;=\; b_k + c_k \sum_{\ell=0}^{n} a_\ell y^\ell \;, \quad k = 0, 1, \ldots, n \;. \tag{13.3}$$

The last equation, for $k = n$, gives

$$a_n y^n \;=\; b_n + c_n \sum_{\ell=0}^{n} a_\ell y^\ell \;,$$

and therefore

$$\begin{aligned} a_n &= c_n a_n , \\ 0 &= c_n a_{n-1} = c_n a_{n-2} = \ldots = c_n a_1 , \\ 0 &= b_n + c_n a_0 . \end{aligned}$$

The equation $a_n(c_n - 1) = 0$ requires that either $a_n = 0$ or $c_n = 1$. If $a_n = 0$, then $P(x)$ is of degree $n-1$. Then repeating the same argument either $a_{n-1} = 0$ or $c_{n-1} = 1$. And this continues until we reach a constant polynomial. We thus conclude that, if we wish to find non-constant polynomials, it must be $c_n = 1$ and $a_n \neq 0$. Then

$$0 = a_{n-1} = a_{n-2} = \ldots = a_1 ,$$

and

$$b_n = -a_0 .$$

Returning to equation (13.3),

$$a_k y^k = b_k + c_k (a_0 + a_n y^n) , \quad k = 0, 1, \ldots, n , \tag{13.4}$$

we see that

$$0 = c_{n-1} = c_{n-2} = \ldots = c_1 = c_0 ,$$

and

$$b_0 = a_0 , \quad b_1 = b_2 = \ldots = b_{n-1} = 0 .$$

All this gives, again, the result that $P(x) = a_n x^n + a_0$ and $q_1(u) = a_0(1 - u^n)$ and $q_2(u) = u^n$. □

Problem 13.10. *Find all polynomials $P(x)$ for which there exists a polynomial of two variables $Q(u, v)$ such that*

$$P(x + y) = Q(x, P(y)) .$$

Solution 1. Any polynomial of degree zero $P(x) = a_0$ satisfies the conditions of the problem. We will thus seek polynomials of degree $n \geq 1$.

Let $Q(u, v)$ be a polynomial of degree n_1 with respect to u and degree n_2 with respect to v. Then $Q(x, P(y))$ is of degree n_1 with respect to x and of degree nn_2 with respect to y. Therefore the given equation requires that $n = n_1$ and $n = nn_2$ or $n_1 = n$ and $n_2 = 1$. We can write the polynomial $Q(u, v)$ in the form

$$Q(u, v) = q_1(u) + q_2(u)\, v ,$$

where the polynomials $q_1(u), q_2(u)$ are of degree n. The given functional equation then takes the form of a Vincze I functional equation (8.13):

$$P(x + y) = q_1(x) + q_2(x)\, P(y) ,$$

with solution

$$P(x) = ax + b , \quad Q(u, v) = a\, u + v . \qquad \square$$

As with the previous problem, we could finish the solution of this problem without any reference to the Vincze II equation:

Solution 2. Let

$$\begin{aligned} P(x) &= \sum_{k=0}^{n} a_k x^k \,, \\ q_1(u) &= \sum_{k=0}^{n} b_k u^k \,, \\ q_2(u) &= \sum_{k=0}^{n} c_k u^k \,. \end{aligned}$$

Then the given equation translates to

$$\sum_{k=0}^{n} a_k\,(x+y)^k = \sum_{k=0}^{n} b_k\,x^k + \sum_{k=0}^{n}\sum_{\ell=0}^{n} c_k a_\ell\,x^k y^\ell \,. \tag{13.5}$$

The right hand side has all terms $x^k y^\ell$ but the left hand side has only terms $x^k y^\ell$ with $k+\ell \le n$. Therefore the coefficients of the terms $x^k y^\ell$ with $k+\ell > n$ in the right hand side must vanish:

$$\begin{aligned} c_n\,a_\ell &= 0\,, \quad \ell = 1,2,\ldots,n\,, \\ c_{n-1}\,a_\ell &= 0\,, \quad \ell = 2,3,\ldots,n\,, \\ &\ldots \\ c_1\,a_\ell &= 0\,, \quad \ell = n\,. \end{aligned}$$

The last equation $c_1 a_n = 0$ requires either $c_1 = 0$ or $a_n = 0$. If $a_n = 0$ then $P(x)$ is of degree $n-1$. Restarting for the new polynomial we would find $c_1 a_{n-1} = 0$ and either $a_{n-1} = 0$ or $c_1 = 0$. Assuming that $a_{n-1} = 0$, $P(x)$ reduces to a polynomial of degree $n-2$. So, in order not to reduce $P(x)$ to a constant polynomial, we must assume $c_1 = 0$ at some stage. So, without loss of generality, we assume that that's the case in the equations written above and formally treat $a_n \neq 0$ in these equations. Therefore

$$c_1 = c_2 = \ldots = c_n = 0\,.$$

Equation (13.5) now becomes:

$$a_0 + a_1(x+y) + \sum_{k=2}^{n} a_k\,(x+y)^k = \sum_{k=0}^{n} b_k\,x^k + c_0 \sum_{\ell=0}^{n} a_\ell\,y^\ell \,. \tag{13.6}$$

Only the left hand side has mixed terms $x^k y^\ell$. Therefore it must be

$$a_2 = a_3 = \ldots = a_n = 0\,.$$

This leaves all terms x^k with $k \geq 2$ unbalanced. Consequently

$$b_2 = b_3 = \ldots = b_n = 0 .$$

Finally, equating the coefficients of the remaining terms in (13.6),

$$\begin{aligned} a_0 &= b_0 + c_0\, a_0 , \\ a_1 &= b_1 , \\ a_1 &= c_0 a_1 . \end{aligned}$$

From the third equation, $a_1(1 - c_0) = 0$. To have a non-constant polynomial, a_1 *must* necessarily be non-zero. Then $c_0 = 1$. The first equation then requires $b_0 = 0$. Therefore we have arrived at the result

$$P(x) = a_1\, x + a_0 , \quad q_1(u) = a_1\, x , \quad q_2(u) = 1 ,$$

as in the previous solution. □

Chapter 14

Conditional Functional Equations

14.1 The Notion of Conditional Equations

Let's assume that we want to solve a modified version of Jensen's functional equation.

Problem 14.1. *Find the continuous solutions $f : [0,1] \to \mathbb{R}$ of the functional equation*

$$f\left(\frac{x+y}{2}\right) = \frac{f(x)+f(y)}{2}\,, \quad \forall x, y \in [0,1]\,.$$

In this modified problem only the domain of the function f has been changed, from $\mathbb{R}$ to the interval [0,1]. As we have said previously, such a change might change the set of solutions. However, one might still want to proceed in the same way we solved the original problem (see page 89). For $y = 0$, the equation gives

$$f\left(\frac{x}{2}\right) = \frac{f(x)+f(0)}{2} = \frac{f(x)+b}{2}\,,$$

where $b = f(0)$. In this relation we substitute $x = y + z$:

$$f\left(\frac{y+z}{2}\right) = \frac{f(y+z)+b}{2}\,.$$

However, the left hand side can be rewritten from the defining relation and thus

$$f(y+z) = f(y) + f(z) - b\,,$$

which can be rewritten in the form

$$g(x+y) = g(x) + g(y)\,,$$

if we define $g(x) = f(x) - b$.The last equation looks deceptively similar to the linear Cauchy equation (5.1) with $x, y \in [0,1]$. However, it is not exactly the same. The equation is not valid for all $x, y \in [0,1]$; it is valid for all $x, y \in [0,1]$ whose sum $x + y$ also belongs in [0,1]. We cannot apply the equation, for example, to $x = y = 1$. Such a functional equation is called a **conditional** (or **restricted**) functional equation.

Quite interestingly, conditional functional equations are not encountered frequently in mathematical competitions (for a recent problem given in IMO 2008 see Problem 20.151). This does not take into account the fact that certain conditions given for functions may be interpreted as conditional functional equations. For example, the condition (13.2) on the polynomial $P(x)$,

$$P(x) = \frac{x}{x+1}, \quad \forall x \in \{0, 1, \ldots, n\},$$

may be thought of as a conditional functional equation since it is valid only for a subset of the domain. This point of view makes the differentiation of equations into regular and conditional a little vague. I personally prefer to avoid using this terminology but the reader must be aware of it since mathematicians have studied many equations that have been labeled as conditional.

14.2 An Example

In this section, we present just a single example of a conditional equation. We will assume that the interested reader will search the literature for additional information and problems.

Problem 14.2 ([5], Problem 81-1)**.** *Let $f : \mathbb{R}^k \to \mathbb{R}$ be a continuous function satisfying the relation*[1]

$$f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}), \quad \textit{whenever } \vec{u} \perp \vec{v}.$$

Show that there exist a constant c and a vector $\vec{a} \in \mathbb{R}^k$ such that

$$f(\vec{u}) = c\,\|\vec{u}\|^2 + \vec{a} \cdot \vec{u},$$

where $\|\vec{u}\|$ is the magnitude of $\vec{u}$.

Solution. The solution given below follows M. St. Vincent [63].

Let

$$f(\vec{u}) = \frac{f(\vec{u}) + f(-\vec{u})}{2} + \frac{f(\vec{u}) - f(-\vec{u})}{2} \equiv f_{even}(\vec{u}) + f_{odd}(\vec{u}).$$

Obviously the even and odd parts must satisfy the same equation. So we shall study them independently.

Case (The even part). Given any $\vec{u}, \vec{v}$ with equal magnitudes $\|\vec{u}\| = \|\vec{v}\|$, we can define

$$\vec{x} = \frac{\vec{u} + \vec{v}}{2}, \quad \vec{y} = \frac{\vec{u} - \vec{v}}{2}.$$

These vectors are perpendicular, i.e. $\vec{x} \cdot \vec{y} = 0$. Also

$$\begin{aligned} f_{even}(\vec{u}) &= f_{even}(\vec{x} + \vec{y}) = f_{even}(\vec{x}) + f_{even}(\vec{y}), \\ f_{even}(\vec{v}) &= f_{even}(\vec{x} + (-\vec{y})) = f_{even}(\vec{x}) + f_{even}(\vec{y}), \end{aligned}$$

[1] The condition $\vec{u} \perp \vec{v}$ is read $\vec{u}$ is perpendicular to $\vec{v}$ and it is equivalent to $\vec{u} \cdot \vec{v} = 0$.

or

$$f_{even}(\vec{u}) = f_{even}(\vec{v}) .$$

The function f_{even} has equal values for vectors of equal lengths. In other words, it cannot depend on the direction of its argument:

$$f_{even}(\vec{u}) = f_{even}(\|\vec{u}\|) .$$

For simplicity we shall indicate $\|\vec{u}\| = u$. In the given functional equation, since $\vec{u} \perp \vec{v}$, $\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 = u^2 + v^2$. Therefore

$$f_{even}\left(\sqrt{u^2 + v^2}\right) = f_{even}(u) + f_{even}(v) , \quad \forall u, v \in \mathbb{R}_+^* .$$

We have thus found an unrestricted equation for the length of the vectors which has been solved in Problem 5.11:

$$f_{even}(u) = c\, u^2 ,$$

with c some constant.

Case (The odd part). First we will establish that

$$f(\lambda\vec{u}) = \lambda\, f(\vec{u}) , \quad \forall\lambda \in \mathbb{R} . \tag{14.1}$$

We shall do so by following the steps of the $\mathbb{NQR}$-method. So, we shall first show that

$$f(n\vec{u}) = n\, f(\vec{u}) , \quad \forall n \in \mathbb{N} . \tag{14.2}$$

Since f_{odd} is an odd continuous function, $f_{odd}(\vec{0}) = 0$. For $n = 1$, the equality is trivially true. Suppose that it is true for some positive integer k. Now, given any $\vec{u} \neq \vec{0}$, consider the vector $\vec{v}$ which is such that $\vec{u} \cdot \vec{v} = 0$ and has length $\|\vec{v}\| = \sqrt{k}\,\|\vec{u}\|$. This vector has the property that

$$(k\vec{u} - \vec{v}) \perp (\vec{u} + \vec{v}) .$$

Therefore from the defining equation

$$f_{odd}(k\vec{u} - \vec{v}) + f_{odd}(\vec{u} + \vec{v}) = f_{odd}((k+1)\vec{u}) .$$

However, each term in the left hand side can be rewritten using that equation too since the arguments are sums of perpendicular vectors. Finally

$$(k+1)\, f_{odd}(\vec{u}) = f_{odd}((k+1)\vec{u}) .$$

That is, equation (14.2) is true for $k+1$ and thus, by induction, valid for all positive integers. In this equation, we now insert $\frac{m}{n}\vec{u}$, $m, n \in \mathbb{N}$, in place of $\vec{u}$ to find:

$$f_{odd}\left(\frac{m}{n}\vec{u}\right) = \frac{m}{n}\, f_{odd}(\vec{u}) .$$

Therefore we have shown that

$$f_{odd}(q\vec{u}) \;=\; q\, f_{odd}(\vec{u})\,, \quad \forall q \in \mathbb{Q}\,.$$

Using the continuity of f_{odd} and a sequence of rational numbers $\{q_n\}$ that converges to a real number λ we finally arrive at advertised result.

Now we shall show that

$$f_{odd}(\vec{u}+\vec{v}) \;=\; f_{odd}(\vec{u}) + f_{odd}(\vec{v})\,, \quad \forall \vec{u}, \vec{v}\,. \tag{14.3}$$

If $\vec{u} = \vec{v} = \vec{0}$ this equation is trivially true. So, without loss of generality, given the vectors $\vec{u}, \vec{v}$ we will assume that $\vec{u} \neq \vec{0}$ and define $\mu = \vec{u}\cdot\vec{v}/||\vec{u}||^2$. Then

$$\vec{u} \perp (\vec{v} - \mu\vec{u})\,.$$

Therefore from the defining equation

$$\begin{aligned} f_{odd}(\vec{u}+\vec{v}) &= f_{odd}((1+\mu)\vec{u}) + f_{odd}(\vec{v}-\mu\vec{u}) \\ &= (1+\mu)\, f_{odd}(\vec{u}) + f_{odd}(\vec{v}-\mu\vec{u}) \\ &= f_{odd}(\vec{u}) + \mu\, f_{odd}(\vec{u}) + f_{odd}(\vec{v}-\mu\vec{u}) \\ &= f_{odd}(\vec{u}) + f_{odd}(\mu\vec{u}+\vec{v}-\mu\vec{u}) \\ &= f_{odd}(\vec{u}) + f_{odd}(\vec{v})\,, \end{aligned}$$

as desired. Equation (14.3) has been solved in Chapter 9 (see Problem 9.1), the solution being

$$f_{odd}(\vec{u}) \;=\; \vec{a}\cdot\vec{u}\,,$$

where $\vec{a}$ some constant vector. □

Chapter 15

Functional Inequalities

15.1 Useful Concepts and Facts

So far we have studied functional relations that contained only equalities between the various terms. Some inequalities did appear but they were not the central piece in any problem except the problem at the end of Section 2.1 (see page 39). In this chapter we study some basic theory and problems that deal with functions satisfying inequalities.

Definition 15.1. A function f is called

- **superadditive** if
$$f(x+y) \geq f(x) + f(y) \ , \quad \forall x, y \ ;$$
- **subadditive** if
$$f(x+y) \leq f(x) + f(y) \ , \quad \forall x, y \ ;$$
- **weakly convex** if
$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2} \ , \quad \forall x, y \ ;$$
- (**strongly**) **convex** if
$$f(tx + (1-t)y) \leq t\, f(x) + (1-t)\, f(y) \ , \quad \forall x, y \ , \quad 0 \leq t \leq 1 \ .$$
- (weakly/strongly) **concave** if $-f$ is (weakly/strongly) convex.

Example 15.1. The function $f(x) = |x|$ with $x \in \mathbb{R}$ is subadditive since $|x+y| \leq |x| + |y|$ while the function $f(x) = x^2$ with $x \in \mathbb{R}_+^*$ is superadditive since $(x+y)^2 = x^2 + y^2 + 2xy \geq x^2 + y^2$. Finally, the functions $f(x) = -\ln x$, $f(x) = 1/x$ defined in $\mathbb{R}_+^*$ are convex and therefore $f(x) = \ln x$, $f(x) = -1/x$ are concave. □

Comment. Convex and concave functions are important means in proving old and new inequalities. However, this seems to me a topic more appropriate for a book in inequalities (see for example [23]). Here I will be brief and present only a quick example.

If f is convex then, from the defining equation, one can show that

$$\frac{\sum_{k=1}^{n} f(x_i)}{n} \geq f\left(\frac{\sum_{k=1}^{n} x_i}{n}\right),$$

for any n and real numbers $x_1, x_2, \ldots, x_n$. This inequality (as its special case for $n = 2$) is known as the **Jensen inequality**.

For the convex function $f(x) = -\ln x$ and n positive numbers $x_1, x_2, \ldots, x_n$, the **Jensen inequality** gives

$$\frac{\sum\limits_{k=1}^{n} x_i}{n} \geq \sqrt[n]{\prod_{k=1}^{n} x_i},$$

which is known as the **Cauchy inequality**. □

The following statements are true.

- Obviously, a convex function is also weakly convex but the inverse is not necessarily true.
- If f is weakly convex and discontinuous at a point, then f is discontinuous everywhere.
- If f is weakly convex and continuous at a point, then f is continuous everywhere.
- If f is weakly convex and continuous, then f is convex.
- If f is convex, then f is continuous.

In the remaining section we discuss a few statements that are less known to students at a lower level.

Theorem 15.1. *If $f'(x)$ exists on $(0, +\infty)$, then $\frac{f(x)}{x}$ is decreasing (increasing) on $(0, +\infty)$ if and only if $f'(x) \leq \frac{f(x)}{x}$ $\left(\text{respectively } f'(x) \geq \frac{f(x)}{x}\right)$ on $(0, +\infty)$.*

Proof. If $f'(x) \leq \frac{f(x)}{x}$, since $x > 0$ we can equivalently write it as

$$\frac{xf'(x) - f(x)}{x^2} \leq 0 \Leftrightarrow \frac{d}{dx}\left(\frac{f(x)}{x}\right) \leq 0,$$

that is, $f(x)/x$ is decreasing (and vice versa). □

Theorem 15.2. *If $\frac{f(x)}{x}$ is decreasing (increasing) on $(0, +\infty)$, then $f(x)$ is subadditive (superadditive).*

Proof. If $f(x)/x$ is decreasing, then

$$f(x+y) = x\,\frac{f(x+y)}{x+y} + y\,\frac{f(x+y)}{x+y} \le x\,\frac{f(x)}{x} + y\,\frac{f(y)}{y} = f(x) + f(y)\,,$$

that is, $f(x)$ is subadditive. □

Obviously the previous two theorems can be combined into the following one that provides a sufficient condition for subadditivity (superadditivity):

Theorem 15.3. *If $f'(x)$ exists and $f'(x) \le \frac{f(x)}{x}$ $\left(\text{respectively } f'(x) \ge \frac{f(x)}{x}\right)$ on $(0, +\infty)$, then $f(x)$ is subadditive (respectively superadditive).*

Another condition used on functions often is the Lipschitz condition. We say that f satisfies the **Lipschitz condition of order α at** x_0 if there exist a constant $M > 0$ (that may depend on x_0) and an interval $(x_0 - s, x_0 + s)$ such that

$$|f(x) - f(x_0)| \le M\,|x - x_0|^{\alpha}\,,$$

for all $x \in (x_0 - s, x_0 + s)$.

Theorem 15.4. *If f satisfies the Lipschitz condition of order α, with $\alpha > 0$, at x_0, then f is continuous at x_0 but not necessarily differentiable.*

Proof. Let $\varepsilon > 0$ be arbitrary, and choose $\delta(\varepsilon) = (\varepsilon/M)^{1/\alpha} > 0$. Without loss of generality, we shall take $\varepsilon < s^{\alpha}M$, where is s is the radius of the neighborhood of the points satisfying the Lipschitz condition. This implies that $\delta(\varepsilon) < s$ and all points x satisfying $|x - x_0| < \delta$ are points in $(x_0 - s, x_0 + s)$. For these points, the Lipschitz condition gives

$$|f(x) - f(x_0)| \le M\,|x - x_0|^{\alpha} \le M\,\delta^{\alpha} = \varepsilon\,,$$

that is, $\lim_{x \to x_0} f(x) = f(x_0)$. □

Theorem 15.5. *If f satisfies the Lipschitz condition of order α, with $\alpha > 1$, at x_0, then f is differentiable at x_0.*

Proof. Let

$$g(x) = \frac{f(x) - f(x_0)}{x - x_0}\,.$$

Obviously $|g(x)| = |f(x) - f(x_0)|/|x - x_0|$ for which we have from the Lipschitz condition

$$|g(x)| \le M\,|x - x_0|^{\beta}\,,$$

with $\beta = \alpha - 1 > 0$. Therefore $g(x)$ is continuous at x_0 from the last result and $f(x)$ is differentiable at x_0. □

15.2 Solved Problems

Problem 15.1. *We say that f satisfies the* ***uniform Lipschitz condition of order*** α *on the interval I if there exists a constant $M > 0$ such that*

$$|f(x) - f(y)| \leq M\,|x - y|^{\alpha}\ ,$$

for all $x, y \in I$.

Show that, if $\alpha > 1$, then f is constant on I.

Solution. The function f is continuous and differentiable at any point of I. Then

$$\left|\frac{f(x) - f(y)}{x - y}\right| \leq M\,|x - y|^{\alpha - 1}\ .$$

In the limit $x \to y$,

$$|f'(y)| \leq 0 \ \Rightarrow\ f'(y) \ = \ 0\ ,$$

or $f(x) = \text{constant}$. □

Problem 15.2 (China 1983; Hungary 1987)**.** *A function f defined on the interval* $[0, 1]$ *satisfies* $f(0) = f(1)$ *and*

$$|f(x_2) - f(x_1)| < |x_2 - x_1|\ , \quad x_1 \neq x_2\ .$$

Prove that

$$|f(x_2) - f(x_1)| < \frac{1}{2}\ , \quad \forall x_1, x_2\ .$$

Solution. There are two possibilities: either $|x_2 - x_1| \leq \frac{1}{2}$ or $|x_2 - x_1| > \frac{1}{2}$. If $|x_2 - x_1| \leq \frac{1}{2}$ then from the given inequality

$$|f(x_2) - f(x_1)| < |x_2 - x_1| \leq \frac{1}{2}\ .$$

If $|x_2 - x_1| > \frac{1}{2}$, without loss of generality we can assume that $x_2 > x_1$. Then $x_2 - x_1 > \frac{1}{2}$ and

$$\begin{aligned}
|f(x_2) - f(x_1)| &= |f(x_2) - f(1) + f(0) - f(x_1)| \\
&\leq |f(x_2) - f(1)| + |f(0) - f(x_1)| \\
&< |x_2 - 1| + |0 - x_1| \\
&= 1 - x_2 + x_1 \ = \ 1 - (x_2 - x_1) \\
&< 1 - \frac{1}{2} = \frac{1}{2}\ .
\end{aligned}$$

□

Problem 15.3 (IMO 1972). *Let f and g be real-valued functions defined for all real values of x and y, and satisfying the equation*

$$f(x+y)+f(x-y) \;=\; 2f(x)g(y)\,,$$

for all x, y. Prove that if $f(x)$ is not identically zero, and if $|f(x)| \leq 1$ for all x, then $|g(y)| \leq 1$ for all y.

Solution. Assume that $|g(y)| \leq 1$ is not true. Then there is a y_0 such that $|g(y_0)| = \lambda > 1$. We shall show that this leads to a contradiction.

Since $f(x)$ is not identically zero, there is also an x_0 such that $f(x_0) \neq 0$. Now consider x_1 to be that point of the set $\{x_0+y_0, x_0-y_0\}$ that gives the greater absolute value for $f(x)$. Then, from the defining equation and the properties of the absolute value

$$2|f(x_1)| \geq |f(x_0+y_0)| + |f(x_0-y_0)| \geq |f(x_0+y_0)+f(x_0-y_0)| \;=\; 2\,|f(x_0)|\,|g(y_0)|\,,$$

or

$$|f(x_1)| \geq \lambda\,\mu\,.$$

where we set $|f(x_0)| = \mu < 1$.

Now let x_2 be that point in the set $\{x_1+y_0, x_1-y_0\}$ that gives the greater absolute value for $f(x)$. As before,

$$2|f(x_2)| \geq |f(x_1+y_0)| + |f(x_1-y_0)| \geq |f(x_1+y_0)+f(x_1-y_0)| \;=\; 2\,|f(x_1)|\,|g(y_0)|\,,$$

or

$$|f(x_2)| \geq \lambda^2\,\mu\,.$$

Continuing in the same fashion, one can define the sequence of points x_n such that

$$|f(x_n)| \geq \lambda^n\,\mu\,.$$

Since $\lambda > 1$, there exists n_0 such that $\lambda^{n_0}\,\mu > 1$, and we have thus proved our claim. □

Problem 15.4 (USA 2000). *Call a real-valued function f **very convex** if*

$$\frac{f(x)+f(y)}{2} \geq f\left(\frac{x+y}{2}\right) + |x-y|$$

holds for all real numbers x and y. Prove that no very convex function exists.

Solution. Let's assume that a very convex function f exists. We will show that this leads to a contradiction.

Without loss of generality we may assume $y \geq x$. We define the sequence of the four points that divide the interval (x, y) in three equal pieces:

$$x_0 = x \;, \quad x_1 = x + \varepsilon \;, \quad x_2 = x + 2\varepsilon \;, \quad x_3 = x + 3\varepsilon = y \;,$$

where, of course, $\varepsilon = (y - x)/3$. We then apply the given functional inequality for the pair of points (x_0, x_2), (x_1, x_3),

$$\begin{aligned} \frac{f(x_0) + f(x_2)}{2} &\geq f(x_1) + 2\varepsilon \;, \\ \frac{f(x_1) + f(x_3)}{2} &\geq f(x_2) + 2\varepsilon \;, \end{aligned}$$

Adding these inequalities we find

$$\frac{f(x_0) + f(x_3)}{2} \geq \frac{f(x_1) + f(x_2)}{2} + 4\varepsilon \;.$$

It then follows that

$$\frac{f(x_0) + f(x_3)}{2} \geq f\left(\frac{x_1 + x_2}{2}\right) + 5\varepsilon \;,$$

or

$$\frac{f(x) + f(y)}{2} \geq f\left(\frac{x + y}{2}\right) + \frac{5}{3}(y - x) \;.$$

Therefore, if the function f satisfies the inequality given with $|x - y|$ in the right hand side, it will also satisfy it with $5|x - y|/3$. By iteration, the function satisfies the inequality

$$\frac{f(x) + f(y)}{2} \geq f\left(\frac{x + y}{2}\right) + \left(\frac{5}{3}\right)^n |x - y| \;,$$

for any n. However the right hand side increases without limit as n increases and therefore the inequality cannot be satisfied. □

Part V

EQUATIONS WITH NO PARAMETERS

Chapter 16

Iterations

Functional equations for functions of a single variable and no 'parameters' involved have appeared in this book in several places. In particular, the majority of the problems in Chapter 4 are such functional equations. In general, given a function $f(x)$, a functional equation satisfied by $f(x)$ is much harder to solve when no 'parameters' enter in the functional equation. The related equations that we have studied so far were simple enough and we did not have to deal with many of the rich and interesting features encountered in the theory of functional equations with no 'parameters'. In this and the following chapter, we will highlight some of these features. However, an exhaustive presentation of the topic is beyond the scope of this book. The interested reader should consult [39] or, at a higher level, Kuczma's book [27] for a thorough and deeper coverage of the topic by a highly-regarded expert.

16.1 The Need for New Methods

Some important functional equations in this category are:

$$\begin{aligned} f^2(x) &= 0\,, & (16.1)\\ f^2(x) &= x\,, & (16.2)\\ f(g(x)) &= f(x) + a\,, \quad a \neq 0\,, & (16.3)\\ f(g(x)) &= \lambda\, f(x)\,, \quad \lambda \neq 0\,, & (16.4)\\ f(g(x)) &= (f(x))^p\,, \quad p \neq 1\,. & (16.5) \end{aligned}$$

Equation (16.3) is known as the **Abel equation**; equation (16.4) is known as the **Schröder equation** and equation (16.5) is known as the **Böttcher equation**.

For $x = y$, the four Cauchy functional equations that we have studied in previous chapters become:

$$\begin{aligned} f(2x) &= 2f(x)\,, \\ f(2x) &= f(x)^2\,, \\ f(x^2) &= 2f(x)\,, \\ f(x^2) &= f(x)^2\,, \end{aligned}$$

special cases of the Schröder (first and third) and Böttcher equations (second and fourth) with the obvious solutions

$$\begin{aligned} f(x) &= a\,x\,, \\ f(x) &= a^x\,, \\ f(x) &= \log_a x\,, \\ f(x) &= x^a\,. \end{aligned}$$

One would like to know under what conditions the above 'no-parameter' Cauchy equations *uniquely* characterize the linear, the exponential, the logarithmic and the power functions, respectively.

Equations (16.1) and (16.2) are special cases of the equation

$$f^n(x) = h(x)\,,$$

where $h(x)$ is a known function and $n > 1$ is a given integer. The solutions $f(x)$ are called the **iterative n-th roots** of h. In particular, the solutions of (16.2) (the iterative square roots of the identity function) are called **involutions**.

Unfortunately, the methods we have presented so far are not adequate to solve functional equations with no 'parameters'. To demonstrate this claim, let's work out the solution of equation (16.1) that consists the simplest possible example.

Problem 16.1. *Let $f : I \to I$ where $I = [a, b]$, $a \le 0 \le b$, be such that $f^2(x) = 0$. Find all such solutions not identically zero.*

Solution. We define

$$\begin{aligned} K &= \{x \in I \mid f(x) = 0\}\,, \\ N &= \{x \in I \mid f(x) \neq 0\}\,. \end{aligned}$$

We shall show that a function f will be a solution if and only if $K \neq \emptyset$, $f(N) \subseteq K$, $0 \in K$.

Necessary: Since $f(x) \not\equiv 0$, there exists an $x_0 \in I$ such that $f(x_0) \neq 0$. Therefore the set N is non-empty.

We set $y_0 = f(x_0)$. Then $f(y_0) = 0$ since $f(f(x_0)) = 0$ by definition. The point y_0 belongs to K which is thus non-empty too.

Similarly, for any $x' \in N$ we can set $y' = f(x')$ and by the defining equation $f(y') = 0$, that is, $y' \in K$. Therefore $f(N) \subseteq K$.

The point 0 must be in K, that is $f(0) = 0$. If this is not true, then for any point $k \in K$, $f(f(k)) = f(0) \neq 0$ which contradicts the defining equation.

Sufficient: Let K and N be two non-empty sets such that $0 \in K$, $K \cap N = \emptyset$ and $K \cup N = I$. Now let $\tilde{f}$ be any function $\tilde{f} : N \to K$. The extension of $\tilde{f}$ to a function $f : I \to I$ by

$$f(x) = \begin{cases} 0\,, & \text{if } x \in K\,, \\ \tilde{f}(x)\,, & \text{if } x \in N\,, \end{cases}$$

is easily verified to satisfy $f^2(x) = 0$ for all $x \in I$. □

Question. *What are the continuous solutions?*

Answer. Since $f(0) = 0$, the continuity guarantees that the function vanishes in a neighborhood of zero. In particular, there $a \leq a' < 0$ and $0 < b' \leq b$ such that $f(x) = 0$ for all $x \in [a', b']$. In the intervals $[a, a')$ and $(b', b]$ the function can take any value (consistent with continuity) in the interval $[a', b']$.

Comment. Notice that $[a', b']$ is not necessarily equal to the set K which may include additional points.

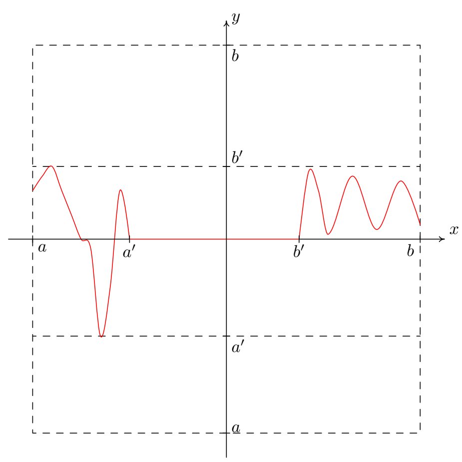

Figure 16.1: The graph of a continuous function that satisfies $f^2(x) = 0$.

In the previous example, we found all solutions of the functional equation. Often, we are interested only in identifying a particular subset of solutions, not necessarily all solutions. This simplifies the task considerably. For example, we can find solutions of the equation (16.2) relatively easily. It is immediate that reflections and inversions are solutions:

$$f_1(x) = -x\,, \quad f_2(x) = \frac{1}{x}\,.$$

However, these are not the only solutions. With a little more effort one can discover that given an invertible function g,

$$f(x) = g^{-1}(-g(x))$$

is an involution. We have thus discovered a family of solutions parameterized by the invertible function g but, of course, we have not shown that these are the only solutions. (For a complete solution see Problem 16.2.)

16.2 Iterates, Orbits, Fixed Points, and Cycles

The iterates of a function $f|X$ have already been defined in section 1.3:

$$f^2 = f \circ f , \quad f^n = f \circ f^{n-1} , \quad n > 2 .$$

For convenience we may often define $f^0 = \mathrm{id}_X$. It is noteworthy that any two iterates of f commute:

$$f^n \circ f^m = f^m \circ f^n = f^{n+m} , \quad \forall n, m \in \mathbb{N} .$$

The sequence of points

$$\{x_n = f^n(x)\}_{n \in \mathbb{N}}$$

is called the **splinter** of x.

Consider two points $x, y \in X$. We call them **equivalent points** under iteration of f and write $x \sim y$ (more exactly $x \sim_f y$ but we will omit the subscript f when there is no ambiguity) if there exist $n, m \in \mathbb{N}$ such that $f^n(x) = f^m(y)$. This relation between points satisfies the three properties of an equivalence relation:

1. *Reflexivity*: for all $x \in X$, $x \sim x$.
2. *Symmetry*: for all $x, y \in X$, if $x \sim y$ then $y \sim x$ too.
3. *Transitivity*: for all $x, y, z \in X$, if $x \sim y$ and $y \sim z$, then $x \sim z$ too.

Therefore $\sim$ is an equivalence relation. We call the corresponding equivalence classes **orbits**. That is, the orbit O_x of a point x is

$$O_x = \{y \in X \mid x \sim y\} .$$

The orbits partition X into a union of pairwise disjoint sets.

If $x_0 \in X$ is such that

$$f^k(x_0) = x_0 , \quad \text{and} \quad f^\ell(x_0) \neq x_0 , \quad \ell = 1, \ldots, k-1 ,$$

then x_0 is called a **fixed point of order** k of f. Fixed points x_* of order 1 are characterized by

$$f(x_*) = x_*$$

and they are simply called **fixed points**. Apparently, a fixed point of order k is a fixed point of the function $g = f^k$.

If $x_0 \in X$ is a fixed point of order k, then all the points in the sequence

$$x_n = f^n(x_0) , \quad n \in \mathbb{N} ,$$

are also fixed points of order k since $(f^n)^k = (f^k)^n$. There are exactly k distinct points in this sequence which form a set called a **cycle of order** k or, simply, a k-**cycle**. For this reason, x_0 is also called a **periodic point** of period k.

Example 16.1. Consider the function $f(x) = x^2 - 2\Big|[-2, 2]$. How many fixed points of order n (if any) does it have? It is actually easy to see that there will be at least one for each n. Looking back at Problem 1.6, we proved that the equation $f^n(x) = x$ has exactly 2^n real roots. Since a fixed point of order n should not be a fixed point of any lower order, and there are

$$2^0 + 2^1 + 2^2 + \cdots + 2^{n-1} = 2^n - 1$$

total roots for $f^k(x) = x$, $k = 1, 2, \ldots, n-1$, the equation $f^n(x) = x$ has at least one root different from all roots of the previous equations. To exactly count the fixed points of order n is a simple exercise in number theory but, unfortunately, requires knowledge of at least one important piece of number theory. For the reader's curiosity and satisfaction, I present the counting at the end of the example. Here, to uncover the pattern, for illustrative purposes, I shall work by example. (Any calculations are automatic from the results of Problem 1.6.)

•The fixed points (of order 1) are

$$x_* = 2\cos 0\,, \quad x_{**} = 2\cos\frac{2\pi}{3}\,.$$

These two points, satisfying $f(x) = x$, will also be satisfying $f^n(x) = x$ for any $n > 1$. Therefore, they must be removed from the solutions of $f^n(x) = x$ which are the candidates for fixed points of order n.

•The roots of $f^2(x) = x$ and thus the candidates for fixed points of order 2 are

$$x_* = 2\cos 0\,, \quad x_{**} = 2\cos\frac{2\pi}{3}\,, \quad a_2 = 2\cos\frac{2\pi}{5}\,, \quad b_2 = 2\cos\frac{4\pi}{5}\,.$$

The first two solutions are those to be excluded. The remaining two solutions however are genuine fixed points of second order. These points, satisfying $f^2(x) = x$, will be satisfying also $x = f^{2k}(x)$, for any $k > 0$. We also notice that they are members of a 2-cycle:

$$\begin{aligned} f(a_2) &= 2\cos\frac{4\pi}{5} = b_2\,, \\ f(b_2) &= 2\cos\frac{8\pi}{5} = 2\cos\left(2\pi - \frac{2\pi}{5}\right) = a_2\,. \end{aligned}$$

In other words, the splinter of any of the two roots, say

$$a_2,\ f(a_2),\ f^2(a_2),\ f^3(a_2),\ f^4(a_2),\ f^5(a_2),\ f^6(a_2),\ f^7(a_2),\ f^8(a_2),\ \ldots$$

has only two distinct points

$$a_2,\ b_2,\ a_2,\ b_2, \ldots$$

•The roots of $f^3(x) = x$ and thus the candidates for fixed points of order 3 are

$$\begin{gathered} x_* = 2\cos 0\,, \quad a_3 = 2\cos\frac{2\pi}{7}\,, \quad b_3 = 2\cos\frac{4\pi}{7}\,, \quad c_3 = 2\cos\frac{6\pi}{7}\,, \\ \alpha_3 = 2\cos\frac{2\pi}{9}\,, \quad \beta_3 = 2\cos\frac{4\pi}{9}\,, \quad x_{**} = 2\cos\frac{2\pi}{3}\,, \quad \gamma_3 = 2\cos\frac{8\pi}{9}\,. \end{gathered}$$

Excluding x_* and x_{**}, the remaining six solutions however are genuine fixed points of third order. These points, satisfying $f^3(x) = x$, will also be satisfying $x = f^{3k}(x)$ for any $k > 0$.

We also notice that a_3, b_3, c_3 and $\alpha_3, \beta_3, \gamma_3$ form two 3-cycles respectively:

$$\begin{aligned} f(a_3) &= 2\cos\frac{4\pi}{7} = b_3 , \\ f(b_3) &= 2\cos\frac{8\pi}{7} = c_3 , \\ f(c_3) &= 2\cos\frac{16\pi}{7} = a_3 , \end{aligned}$$

and

$$\begin{aligned} f(\alpha_3) &= 2\cos\frac{4\pi}{9} = \beta_3 , \\ f(\beta_3) &= 2\cos\frac{8\pi}{9} = \gamma_3 , \\ f(\gamma_3) &= 2\cos\frac{16\pi}{9} = \alpha_3 . \end{aligned}$$

•There are 16 candidates for fixed points of order 4. From them we must remove the fixed points of order 1 (two points) and order 2 (another two points). Therefore, we are left with 12 fixed points of order 4,

$$\begin{aligned} a_4 &= 2\cos\frac{2\pi}{15} , \quad b_4 = 2\cos\frac{4\pi}{15} , \quad c_4 = 2\cos\frac{8\pi}{15} , \quad d_4 = 2\cos\frac{14\pi}{15} , \\ \alpha_4 &= 2\cos\frac{2\pi}{17} , \quad \beta_4 = 2\cos\frac{4\pi}{17} , \quad \gamma_4 = 2\cos\frac{8\pi}{17} , \quad \delta_4 = 2\cos\frac{16\pi}{17} , \\ \aleph_4 &= 2\cos\frac{6\pi}{17} , \quad \beth_4 = 2\cos\frac{12\pi}{17} , \quad \gimel_4 = 2\cos\frac{10\pi}{17} , \quad \daleth_4 = 2\cos\frac{14\pi}{17} . \end{aligned}$$

which form three 4-cycles. We have labeled the points of these cycles by latin, greek, hebrew letters respectively.

The general pattern should be obvious now. For exact counting of the fixed points of order n see the following paragraph that uses a result from number theory.

Counting the n-periodic points: Consider the arithmetic functions f and g defined on $\mathbb{N}$ and assume that they satisfy the relation

$$g(n) = \sum_{d|n} f(d) ,$$

for $n \geq 1$. The notation $d|n$ stands for 'd divides n', that is the sum is over all divisors d of n. This relation can be solved for $f(n)$. The result is

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) ,$$

where $\mu(n)$ is the Möbius function defined in Example 3.3. This formula is known as the **Möbius inversion formula**. You may find the proof in any text of number theory, such as [12]. Also, inverting the latter formula gives the former.

Now let's consider any function $f(x)$ for which the equation $f^n(x) = x$ has 2^n solutions. A fixed point of order d satisfing $x = f^d(x)$ will also be a solution of $x = f^{qd}(x)$ for any an integer q. From this we see that, if $d|n$, then the solutions of $x = f^d(x)$ are also solutions of $x = f^n(x)$. Let $N(d)$ stand for the number of fixed points of order d. Then our discussion implies that

$$\sum_{d|n} N(d) \;=\; 2^n \;,$$

which we can invert to find

$$N(n) \;=\; \sum_{d|n} \mu\left(\frac{n}{d}\right) 2^d \;.$$

For example, if $n = 4$, d is $1, 2, 4$ and thus

$$N(4) \;\;=\;\; \mu(4)\,2 + \mu(2)\,2^2 + \mu(1)\,2^4 \;=\; 0 \cdot 2 - 1 \cdot 4 + 1 \cdot 16 \;=\; 12\,,$$

as found previously. □

16.3 Fixed Points: Discussion

Consider a function f and a point x_0 in its domain. If x_0 is a fixed point, then its splinter is a constant sequence $x_0, x_0, x_0, \cdots$. But, what happens if x_0 is not a fixed point? To answer this question, we will first demonstrate how to construct the splinter graphically. This is best done in Figure 16.2 which the reader should consult to understand the construction.

For the graph Γ given in Figure 16.2, we see that the splinter of a point x_0 (less than the fixed point x_*) creates an increasing sequence of points which converges to x_*. In fact we can easily prove the following theorem:

Theorem 16.1. *If f is continuous and the splinter converges to a limit ℓ, then ℓ must be a fixed point.*

Proof. Since any subsequence of a converging sequence converges to the same limit, we have:

$$\begin{aligned} \ell &= \lim_{n\to\infty} f^n(x_0) \\ &= \lim_{n\to\infty} f^{n+1}(x_0) \\ &= \lim_{n\to\infty} f(f^n(x_0)) \\ &= f(\lim_{n\to\infty} f^n(x_0)) \\ &= f(\ell)\,. \end{aligned}$$

□

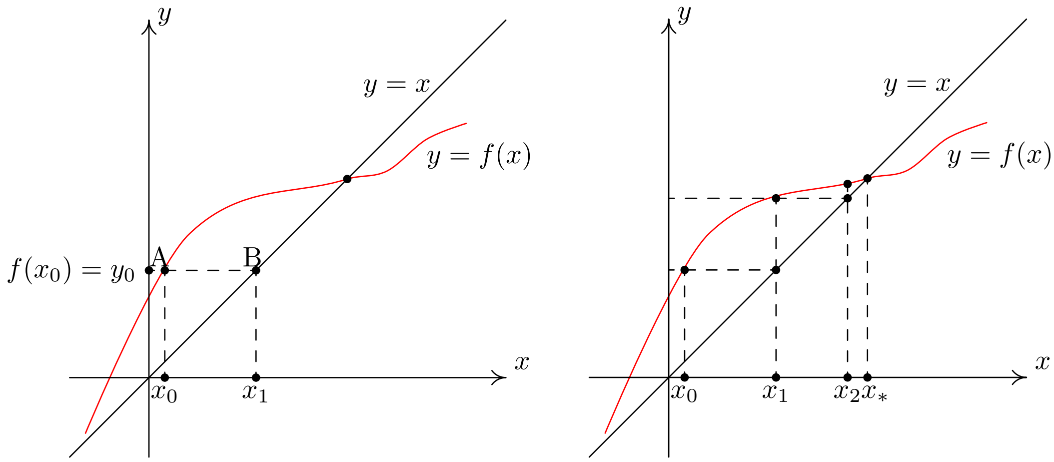

Figure 16.2: Left: Graphical representation of the map $x_0 \mapsto x_1 = f(x_0)$. First draw the graph Γ of the function $y = f(x)$ and the diagonal $y = x$. Given the point x_0 we raise a vertical line to find the intersection point A with the graph Γ. Point A has coordinates $(x_0, f(x_0))$. We then draw a horizontal line through A that intersects the diagonal at point B with coordinates $(f(x_0), f(x_0))$. Projecting this point on the x-axis determines the point x_1 with value $f(x_0)$. Right: Graphical construction of the splinter of x_0. Using the procedure described in the left picture, we can construct the successive images $x_0 \mapsto x_1 \mapsto x_2 \mapsto \cdots$. In the particular case shown in this picture, the splinter of x_0 converges to the fixed point x_*.

Figure 16.3 illustrates two different, very special but important cases of fixed points. In the left case the splinter of any point converges to x_*. In the right case, the splinter of any point moves further away from x_*. This motivates us to introduce the definition:

Definition 16.1. A fixed point x_* of a function f is called

(a) **attractive** or **stable** if there is a neighborhood $N(x_*)$ of x_* such that for any $x \in N(x_*)$ the splinter of x converges to x_*.

(b) **repulsive** or **unstable** if there is a neighborhood $N(x_*)$ of x_* such that for any $x \in N(x_*) \setminus \{x_*\}$ the splinter of x contains a point not in $N(x_*)$.

Consider the set F of all attractive and repulsive fixed points of a function f. Then each such point is isolated in F, that is, for each $x_* \in F$ there is a neighborhood $N(x_*)$ such that $x'_* \notin N(x_*)$ for all $x'_* \in F \setminus \{x_*\}$. Otherwise, if there was at least one $x'_* \in N(x_*)$, then its splinter would be the constant sequence $x'_*, x'_*, x'_*, \ldots$ which neither converges to x_* nor has a point outside $N(x_*)$.

We can now establish simple criteria that allow us check for fixed points of a function f and their nature.

Theorem 16.2. *Let f be a continuous function such that $f(a) > a$ and $f(b) < b$ at some points a and b. Then f has a fixed point x_* between a and b.*

Proof. Without loss of generality we assume $a < b$ (otherwise we rename the points). Then for the continuous function $g(x) = f(x) - x$ we have $g(a) > 0$ and $g(b) < 0$. Therefore, there must be a point $x_* \in (a, b)$ such that $g(x_*) = 0$ or $f(x_*) = x_*$. $\square$

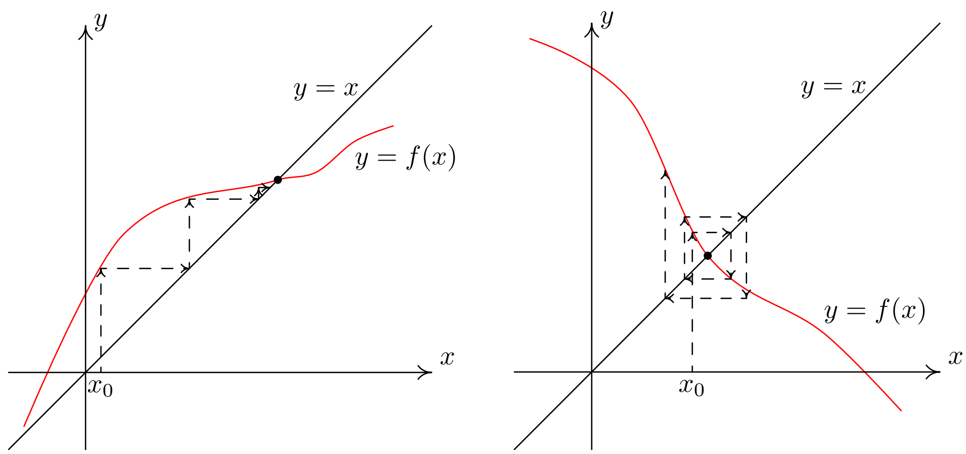

Figure 16.3: Left: An attractive fixed point. Right: A repulsive fixed point.

Corollary (Brouwer Fixed Point Theorem). *Suppose that $f : [a, b] \to [a, b]$ is a continuous map. Then there exists a fixed point $x_* \in [a, b]$.*

Proof. If $f(a) = a$ or $f(b) = b$, we are done. If that's not the case, then it must be that $f(a) > a$ and $f(b) < b$ and from the previous theorem there is an $x_* \in (a, b)$ such that $f(x_*) = x_*$. □

Question. *The function $f : [1, +\infty) \to [1, +\infty)$ with*

$$x \mapsto f(x) = x + \frac{1}{x}$$

has no fixed point $x_* \in (1, +\infty)$. How is this possible? Has any implicit assumption been introduced in the previous proof?

Answer. There has been no implicit assumption in the proof of Brouwer fixed point theorem. The theorem has been proved for an interval $[a, b]$ that has the following properties: it is *closed*, *bounded* and *connected* (that is, not the union of two disjoints intervals). If any of these properties is removed, then the statement may not be valid.

Theorem 16.3. *Let f be a continuous function and x_* be a fixed point of f. Then, if for all points x in some neighborhood $N(x_*)$ of x_* we have*

$$\left|\frac{f(x) - f(x_*)}{x - x_*}\right| < 1 \quad \Rightarrow \quad x_* \text{ is an attractive fixed point;}$$
$$\left|\frac{f(x) - f(x_*)}{x - x_*}\right| > 1 \quad \Rightarrow \quad x_* \text{ is a repulsive fixed point.}$$

Proof. We will prove the first statement; the second statement is proved similarly with the appropriate changes.

In the relation $|f(x) - f(x_*)| < |x - x_*|$ we replace x by the points defined by the splinter of x: $x_n = f^n(x)$. Then

$$|f(x_n) - f(x_*)| < |x_n - x_*| \;\Rightarrow\; |x_{n+1} - x_*| < |x_n - x_*| \,,$$

for any $n > 1$. Therefore the sequence

$$0 < \cdots < |x_2 - x_*| < |x_1 - x_*| < |x - x_*|$$

is bounded and decreasing and hence converges. As such, the splinter converges, too.

Let

$$\left| \frac{f(x) - x_*}{x - x_*} \right| \;=\; \lambda < 1 \,.$$

replacing x with $f(x)$, we find:

$$\left| f^2(x) - x_* \right| \;=\; \lambda \left| f(x) - x_* \right| \;=\; \lambda^2 \,|x - x_*| \,.$$

Inductively,

$$\left| f^n(x) - x_* \right| \;=\; \lambda^n \,|x - x_*| \,.$$

In the limit $x \to x_*$, $f^n(x) \to x_*$. □

Corollary. *Let f be a continuous function and x_* be a fixed point of f. If the derivative $f'(x_*)$ exists, then if*

$$\begin{aligned} \left| f'(x_*) \right| < 1 \quad &\Rightarrow \quad x_* \textit{ is an attractive fixed point;} \\ \left| f'(x_*) \right| > 1 \quad &\Rightarrow \quad x_* \textit{ is a repulsive fixed point.} \end{aligned}$$

Proof. The corollary is immediate from the last theorem if we take the limit $x \to x_*$. □

Obviously, if $|f'(x_*)| = 1$ — we call x_* a **marginal point** — we need to look at higher order derivatives to decide about its behavior. Also, if $f'(x_*) = 0$, we call x_* **superstable**.

A final comment is in order here: attractive and repulsive fixed points are very special. Look, for example, at the two functions

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

For the former all points are fixed points while for the latter $x_* = 0$ is the unique fixed point. However, none of these fixed points is attractive or repulsive. The splinter of any point in the neighborhood of a fixed point neither converges nor moves away from the fixed point.

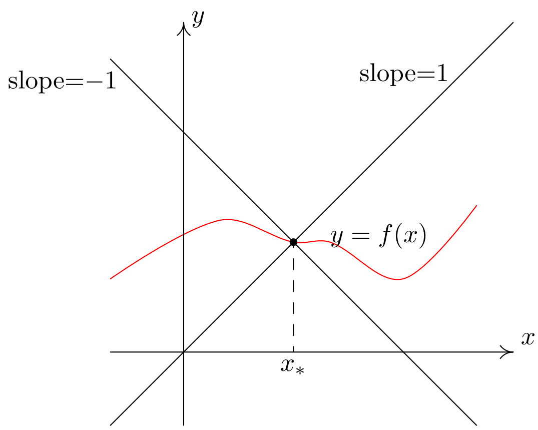

Figure 16.4: Geometric interpretation of the criteria that characterize the nature (attractive vs. repulsive) of a fixed point. Notice that $\frac{f(x)-f(x_*)}{x-x_*}$ is the slope of the secant line that joins the fixed point (x_*, x_*) with the point $(x, f(x))$ in a neighborhood of the fixed point. Therefore, for an attractive fixed point all such secants of the graph of $f(x)$ have slopes in the interval $(-1, 1)$ (left picture) while for a repulsive fixed point all such secants have slopes in $(-\infty, -1) \cup (1, +\infty)$ (right picture). A similar statement is true for the slope of the tangent at the fixed point (if the derivative at x_* exists).

16.4 Cycles: Discussion

Consider again the function $f(x) = -x$, with which we closed the previous section. The splinter of any point $x_0 \neq 0$,

$$x_0,\ -x_0,\ x_0,\ -x_0,\ x_0,\ -x_0,\ \dots\ ,$$

is a periodic sequence with period 2. The points $x_0, -x_0$ form a 2-cycle (cycle of order 2). The function

$$f(x) = \begin{cases} 1\,, & \text{if } x = 0\,, \\ 2\,, & \text{if } x = 1\,, \\ 3\,, & \text{if } x = 2\,, \\ \dots & \\ N-1\,, & \text{if } x = N-2\,, \\ 0\,, & \text{otherwise}\,, \end{cases}$$

has a unique N-cycle for any value of $N \in \mathbb{N} \setminus \{0, 1\}$:

$$0,\ 1,\ 2,\ \dots,\ N-1$$

This function is discontinuous. Constructing a continuous function with an N-cycle is not a trivial matter, as there are rules governing the existence of cycles for such functions. We will present these rules with no attempt to prove them, as the corresponding theory is quite complicated and spans a vast area of mathematics.

In the set of natural numbers $\mathbb{N}^*$ consider the following total order

$$\begin{gathered} 3 \prec 5 \prec 7 \prec \cdots \prec 2 \cdot 3 \prec 2 \cdot 5 \prec 2 \cdot 7 \prec \cdots \prec 2^2 \cdot 3 \prec 2^2 \cdot 5 \prec 2^2 \cdot 7 \prec \ldots \\ \cdots \prec 2^n \cdot 3 \prec 2^n \cdot 5 \prec 2^n \cdot 7 \prec \cdots \prec 2^m \prec 2^{m-1} \prec 2^n \cdots \prec 8 \prec 4 \prec 2 \prec 1 \end{gathered}$$

which is known as **Šarkovskii's order**. In this order first come all odd numbers (in the usual order), then come their multiples by 2, then come their multiples by $4 = 2^2$, then come their multiples by $8 = 2^3$, and so on, and finally come the powers of two in decreasing order.

The following theorem is considered by many as one of the most important theorems in modern mathematics[1]:

Theorem 16.4 (Šarkovskii)**.** *Let $f : I \to I$ be a continuous function defined on the interval I. If f has an m-cycle, then it has an n-cycle for any n such that $m \prec n$.*

According to this theorem if f has a 2-cycle, it does not necessarily have any other cycle (except fixed points) — e.g. the function $f(x) = -x$. However, if the function has a 3-cycle, it must have a cycle of every order!

Example 16.2. Let's return to Example 16.1. There we showed that the function $f(x) = x^2 - 2$ on $[-2, 2]$ has 3-cycles. Therefore, according to Šarkovskii's theorem it must have have a cycle of every order. Can you demonstrate it? □

In the previous section, we showed that if the splinter of a continuous function converges, then it must converge to a fixed point. This established only a necessary condition, not a sufficient one. The following theorem improves on this situation.

Theorem 16.5. *Let $f : I \to I$ be a continuous function. The splinter of any $x \in I$ converges to some fixed point if and only if f has only fixed points and no other cycles.*

The next theorem extends the result to a function with specific cycles.

Theorem 16.6. *Let $f : I \to I$ be a continuous function with cycles of order 1, 2, $2^2, \ldots, 2^n$ only. The splinter of any point $x \in I$ converges to some fixed point or cycle.*

The last theorem suggests that, as in the case of fixed points, there are special cycles to which a splinter might converge (attract) or, perhaps, repel. So we introduce the following definition.

Definition 16.2. Let $f : I \to I$ be a differentiable function[2] and $a_1, a_2, a_3, \ldots, a_k$ be a k-cycle of f. Then this cycle is called

(a) **attractive** or **stable** if at least one point of the cycle is an attractive fixed point of f^k;

(b) **repulsive** or **unstable** if at least one point of the cycle is a repulsive fixed point of f^k.

[1] Yet, Šarkovskii's results [56, 57] went unnoticed for many years. The fact that the original arguments by Šarkovskii are quite formidable and that he published his work in Russian did not help. A simplified proof was found by Štefan in 1977 and was published in English [61].

[2] I adopt differentiability to simplify somewhat my discussion and be able to present simple proofs for the theorems that follow.

Lemma 16.1. *If $x_0 \in I$, let*[3]

$$D = \frac{d}{dx} f^n(x)\Big|_{x=x_0} .$$

If $x_1 = f(x_0)$, $x_2 = f(x_1)$, $\ldots$, $x_{n-1} = f(x_{n-2})$, then

$$D = f'(x_0) f'(x_1) \ldots f'(x_{n-1}) .$$

Proof. For $n = 2$, the statement is true by the well known theorem on the differentiation of a composite function,

$$\frac{d}{dx} f(g(x)) = f'(g)\, g'(x) , \tag{16.6}$$

and setting $g = f$, $x = x_0$. Now, let it be true for $n = k$:

$$\frac{d}{dx} f^k(x)\Big|_{x=x_0} = f'(x_0)\, f'(x_1) \ldots f'(x_{k-1}) .$$

We will show that it is also true for $n = k + 1$. In equation (16.6), we set $g = f^k$:

$$\frac{d}{dx} f^{k+1}(x) = f'(f^k(x))\, \frac{d}{dx} f^k(x) .$$

For $x = x_0$ the last equation gives

$$\frac{d}{dx} f^{k+1}(x)\Big|_{x=x_0} = f'(x_0)\, f'(x_1) \ldots f'(x_{k-1}) f'(x_k) .$$

and this completes the proof. □

The next two theorems are useful criteria to identify attractive and repulsive cycles.

Theorem 16.7. *Let f be a differentiable function. Every point of an attractive (repulsive) k-cycle a_1, a_2, a_3, $\ldots$, a_k of f is an attractive (repulsive) fixed point of f^k.*

Proof. Let $g = f^k$. Then from the previous lemma

$$g'(a_1) = g'(a_2) = \ldots = g'(a_k) ,$$

since each derivative is equal to $f'(a_1) f'(a_2) \ldots f'(a_k)$. This requires that all points of an attractive (repulsive) cycle are simultaneously attractive (repulsive). □

Theorem 16.8. *Let $f : I \to I$ be a differentiable function and a_1, a_2, a_3, $\ldots$, a_k be a k-cycle of f. If $D = f'(a_1) f'(a_2) \ldots f'(a_k)$, then*

$$|D| < 1 \quad \Rightarrow \quad \textit{the cycle is attractive},$$
$$|D| > 1 \quad \Rightarrow \quad \textit{the cycle is repulsive}.$$

Proof. Immediate from the previous theorem. □

[3] I have used the symbol 'D' as a shorthand for *discriminant*. A discriminant, as that of a quadratic polynomial, is an instrument that allows us to differentiate between two or more choices.

16.5 From Iterations to Difference Equations

Functional equations for a function $f(x)$ that contains $f : \mathbb{R} \to \mathbb{R}$ and some of its iterations are intimately related to difference equations. Of course, this relation is useful when the corresponding difference equation can be solved more easily than the original functional equation.

To illustrate the idea, consider the functional equation

$$f^3(x) - f^2(x) - 8f(x) \;=\; -12x\,.$$

The point x can be arbitrary, so we can take it to be $x = f^n(y)$:

$$f^{n+3}(y) - f^{n+2}(y) - 8f^{n+1}(y) \;=\; -12f^n(y)\,.$$

Now consider the splinter of y,

$$y, \quad f(y), \quad f^2(y), \quad \ldots, f^n(y), \quad f^{n+1}(y), \quad f^{n+2}(y), \ldots\,,$$

which we can identify as a sequence

$$a_0, \quad a_1, \quad a_2, \quad \ldots, a_n, \quad a_{n+1}, \quad a_{n+2}, \cdots\,.$$

Then, the given functional equation becomes a difference equation (one for every point y)

$$a_{n+3} - a_{n+2} - 8a_{n+1} \;=\; -12a_n\,,$$

with the fundamental set of solutions 2^n, $n2^n$, $(-3)^n$. Therefore, the sequences

$$\begin{aligned} a_n &= A\,2^n\,, \\ a_n &= C\,(-3)^n\,, \end{aligned}$$

solve the difference equation from which we recover the two solutions

$$\begin{aligned} y \;=\; a_0 &= A\,, \quad a_1 \;=\; f(y) \;=\; A\,2\,, \\ y \;=\; a_0 &= C\,, \quad a_1 \;=\; f(y) \;=\; C\,(-3)\,, \end{aligned}$$

that is

$$f(y) \;=\; 2y\,, \quad \text{or} \quad f(y) \;=\; -3y\,.$$

for the original functional equation. The issue of uniqueness remains open unless additional conditions are specified. (See the solved Problem 16.5 and the unsolved Problem 20.135.)

It should be obvious that the reverse method is also useful: starting from a difference equation, one might consider iterates of a function to draw conclusions. (See, for instance, the solved Problem 16.4.)

16.6 Solved Problems

Problem 16.2. *Find all functions $f : I \to I$ with $I = (a, b)$, $a \leq 0 \leq b$ that satisfy equation (16.2).*

Solution. The function f is injective. Indeed, if two images are equal, say $y_1 = y_2$, then $f(x_1) = f(x_2) \Rightarrow f^2(x_1) = f^2(x_2) \Rightarrow x_1 = x_2$.

The defining functional equation implies that all points of I are fixed points of f^2; that is, they are fixed points of order 1 or 2. Therefore either $f(x) = x$ or $f(x) \neq x$ and $f^2(x) = x$. Let's define the sets

$$\begin{aligned} F &= \{x \in I \mid f(x) = x\} , \\ L &= \{x \in I \mid f(x) < x\} , \\ G &= \{x \in I \mid f(x) > x\} . \end{aligned}$$

These are obviously pairwise disjoint and their union equals I. All fixed points belong to F. All remaining points form 2-cycles $(x, f(x))$ and they belong to $L \cup G$. The two points of a cycle (ℓ, ℓ') belong to different sets since, if they were in the same set (say L) there would be a contradiction ($\ell' = f(\ell) < \ell$ and $\ell = f(\ell') < \ell'$.) Therefore, the 2-cycles define a bijective map between L and G. If f_L is the restriction of f to L, $f_L : L \to G$, and f_G is the restriction of f to G, $f_G : G \to L$, then $f_L = f_G^{-1}$.

Inversely, given a partition of I to three sets F, L, G and $g : L \to G$ a bijective function, then the function

$$f(x) = \begin{cases} g(x) , & \text{if } x \in L , \\ x , & \text{if } x \in F , \\ g^{-1}(x) , & \text{if } x \in G , \end{cases}$$

satisfies equation (16.2). □

Problem 16.3 (Putnam 1952)**.** *Given any real number N_0, if $N_{j+1} = \cos N_j$, prove that $\lim_{j\to\infty} N_j$ exists and it is independent of N_0.*

Solution. The sequence of points $\{N_j\}$ is the splinter of the cosine function $f(x) = \cos x$ generated by the point N_0. The cosine function has a fixed point x_* defined by $\cos x_* = x_*$. At the fixed point $f'(x_*) = -\sin x_*$ and thus $|f'(x_*)| < 1$, that is x_* is an attractive fixed point which implies that $\lim_{j\to\infty} N_j$ exists and it is independent of N_0. □

Problem 16.4 (Putnam 1947)**.** *If $\{a_n\}$ is a sequence of numbers such that for $n \geq 1$*

$$(2 - a_n)\, a_{n+1} = 1 ,$$

prove that $\lim a_n$ *as* $n \to \infty$ *exists and is equal to one.*

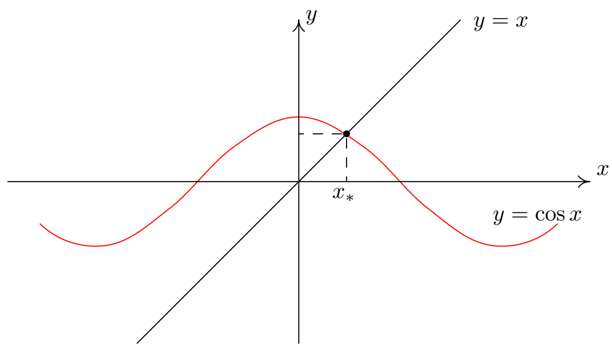

Figure 16.5: The cosine function has a unique fixed point x_* whose approximate value is 0.739.

Solution. The sequence defined by the given recursion relation

$$a_{n+1} = \frac{1}{2 - a_n}$$

can be written as the splinter[4] of the function

$$f(x) = \frac{1}{2 - x}$$

generated by the point $x = a_1$. As in the previous problem we expect that $f(x)$ has an attractive fixed point and this is $x_* = 1$.

Indeed, we can easily confirm that $f(x)$ has a fixed point

$$\frac{1}{2 - x_*} = x_* \Rightarrow (x_* - 1)^2 = 0 \Rightarrow x_* = 1 .$$

However, the criteria to confirm that it is an attractive point fail: At $x_* = 1$, $f'(x_*) = 1$ and the ratio $|f(x) - 1|/|x - 1|$ is less than 1 only on one side of x_*. So, to establish that the splinter converges to 1, additional work is required.

If $x = 1$ then the splinter is the constant sequence $1, 1, 1, \ldots$, that is obviously convergent to 1. If $a_1 \neq 1$ we must consider three cases.

Case I: $x < 1$. When $x < 1$ then

$$\frac{|f(x) - 1|}{|x - 1|} = \frac{1}{|2 - x|} < 1 ,$$

and therefore $f^n(x) \to 1$ as $x \to 1^-$.

[4]This problem may be solved more easily by simpler techniques. However, the use of iterations helps us to demonstrate two ideas: (a) how we can employ a function when a sequence is given and (b) how to modify the techniques presented in theory to verify an attractive fixed point when the theorems we have proved fail.

Case II: $x > 2$. In this case $f(x) < 0$ and the first point in the splinter of x goes back to the previous case.

Case III: $2 > x > 1$. In this case

$$(x-1)^2 > 0 \;\Rightarrow\; 1 > x(2-x) \;\Rightarrow\; \frac{1}{2-x} > x\,,$$

that is, the splinter of x is an increasing sequence. As such at some value of n, the iteration will produce a point $a_n > 2$ which takes us back to case II. (One can check that $f(x) > 2$ when $x > 3/2$.) □

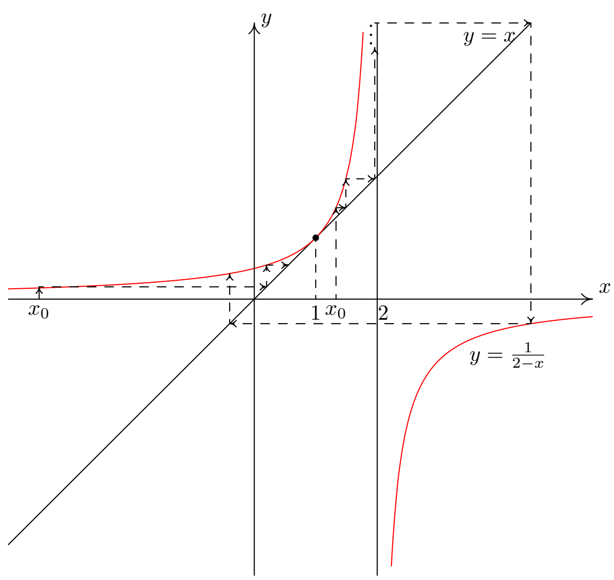

Figure 16.6: The graph of $f(x) = \frac{1}{2-x}$. Shown is also part of the splinter of a point x_0 in the two cases: $x_0 < 1$ and $1 < x_0 < 2$.

Problem 16.5 (Putnam 1988)**.** *Prove that there exists a unique function f from the set $\mathbb{R}_+^*$ of positive real numbers to $\mathbb{R}_+^*$ such that*

$$f(f(x)) \;=\; 6x - f(x)$$

and $f(x) > 0$ for all $x > 0$.

Solution. Setting $f^{n-1}(x)$ in place of x and using the notation $a_n = f^n(x)$, $n = 0, 1, 2, \ldots$, we find the recursion relation

$$a_{n+1} + a_n - 6a_{n-1} \;=\; 0\,.$$

We will seek solutions $a_n = \lambda^n$ which leads to the characteristic equation

$$\lambda^2 + \lambda - 6 = 0 ,$$

with solutions $\lambda_\pm = 2, -3$. Then, 2^n and $(-3)^n$ give a fundamental set of solutions. Since the function sought must be positive, only the first solution is acceptable, and thus $f(x) = 2x$. □

Problem 16.6 (Putnam 1979). *Establish necessary and sufficient conditions on the constant k for the existence of a continuous real-valued function $f(x)$ satisfying*

$$f(f(x)) = kx^9$$

for all real x.

Comment. Notice that the functions f sought are the iterative square roots of kx^9. □

Solution. If $k > 0$ we can see immediately that $f(x) = k^{1/4}\, x^3$ is a solution. Therefore $k > 0$ is a sufficient condition to produce a solution.

To see that this condition is also necessary, we notice that if

$$f(x) = f(y) \Rightarrow f^2(x) = f^2(y) \Rightarrow kx^9 = ky^9 \Rightarrow x = y .$$

The function f must be injective. Since it is also continuous it must be strictly increasing or strictly decreasing and its second iterate f^2 must be strictly increasing. The function kx^9 will be strictly increasing if $k > 0$. □

Problem 16.7 (IMO 1976). *Let $f(x) = x^2 - 2$ and $f^j(x) = f(f^{j-1}(x))$ for $j = 2, 3, \cdots$. Show that for any positive integer n, the roots of the equation $f^n(x) = x$ are real and distinct.*

Solution. If $f(x)$ is a polynomial of degree 2, then $f^n(x)$ is a polynomial of degree 2^n. Therefore, we must prove that $f^n(x)$ has 2^n real and distinct fixed points (not necessarily of order n). Actually, since f is even, f^n is easily seen to be an even polynomial and we can restrict ourselves to the interval $(-\infty, 0]$ since a similar reasoning applies for the interval $[0, +\infty)$. We thus need to prove that f^n has 2^{n-1} negative fixed points[5].

Now we show that, if $x > 2$, then $f^n(x) > x$. For $n = 1$, if $x > 2$ it is easy to see that $f(x) = x^2 - 2 > x^2 - x = x(x-1) > x$. Let it be true for $n = k$. That is, let $f^k(x) > x$ for $x > 2$. Then set $y = f(x)$, with $x > 2$. Obviously $y = f(x) > 2$ and therefore $f^k(y) > y$. But

[5] Have I implied that if x_* is a fixed point, then $-x_*$ is also a fixed point? Answer yes or no and provide an explanation. As a specific example, look at the point $x = 2$.

the last inequality is equivalent to $f^{k+1}(x) > f(x)$, and since $f(x) > x$, also $f^{k+1}(x) > x$. This concludes the proof. By symmetry, $f^n(x) > x$ for $x < -2$.

Now, let's show that f has two fixed points. Of course, solving the quadratic equation $f(x) = x$ we can immediately find $x = 2$ or $x = -1$, that is one positive and one negative fixed point. However, I am going to use an alternative procedure as an example of how to prove the similar statement for the iterations of f. Since $f(a) > a$ for any $a < -2$ and $f(0) = -2 < 0$, by Theorem 16.2, f has a fixed point in $(a, 0)$. Therefore, f has $1 = 2^0$ negative fixed points.

Since $f(-2)f(0) < 0$, by Bolzano's theorem, there is a point a_1 in $(-2, 0)$ such that $f(a_1) = 0$. In the next iteration, $f^2(a_1) = f(0) = -2$. Also, $f^2(0) = f(-2) = 2$. Therefore, the graph of $f^2(x)$ looks as shown in Figure 16.7 (middle plot). In each of the two intervals (a, a_1) and $(a_1, 0)$, f^2 has a fixed point by the use of Theorem 16.2. Therefore, it has 2^1 fixed points in $(-\infty, 0)$.

By Bolzano's theorem, f^2 has also one root in each of the two intervals (a, a_1) and $(a_1, 0)$. Let them be a_2 and a_3. In the next iteration $f^3(a_2) = f(0) = -2$ and similarly for a_3. Also, $f^3(0) = f(2) = 2$. Therefore, the graph of $f^3(x)$ looks as shown in Figure 16.7 (bottom plot). In each of the intervals (a, a_2), (a_2, a_1), (a_1, a_3) and $(a_3, 0)$, f^3 has a fixed point by the use of Theorem 16.2. Therefore, it has 2^2 negative fixed points.

We can continue inductively to find that $f^n(x)$ achieves the value -2 at 2^{n-2} points and the value $+2$ at $2^{n-2} + 1$ points which split the domain $[+2, 0]$ in 2^{n-1} subintervals[6] in each of which the function f^n has a fixed point. □

Having understood the previous solution, I propose for you the following variation of 1998 Turkish Olympiad Problem 1.10:

Problem 16.8. *Consider the function $f : [0, 1] \to [0, 1]$ defined by*

$$f(x) = 4x(1 - x) .$$

How many distinct roots does the equation $f^{1992}(x) = x$ have?

Although, it may be obvious to you now, Figure 16.8 shows the first iterations of $f(x)$. After you have tried it (and hopefully solved it) read Section 16.7.

Problem 16.9 (Putnam 1971)**.** *Determine all polynomials $P(x)$ such that*

$$P(x^2 + 1) = P(x)^2 + 1 ,$$

and $P(0) = 0$.

[6] Although we did not say explicitly, we used the fact that the splinters of ± 2 are $2 \mapsto f(2) = 2 \mapsto 2 \mapsto 2 \mapsto \dots$ and $-2 \mapsto f(-2) = 2 \mapsto 2 \mapsto 2 \mapsto \dots$. Also, if a is a root of f^k, then its splinter — starting at the k-th iteration — will be $a \mapsto 0 \mapsto -2 \mapsto 2 \mapsto 2 \mapsto \dots$.

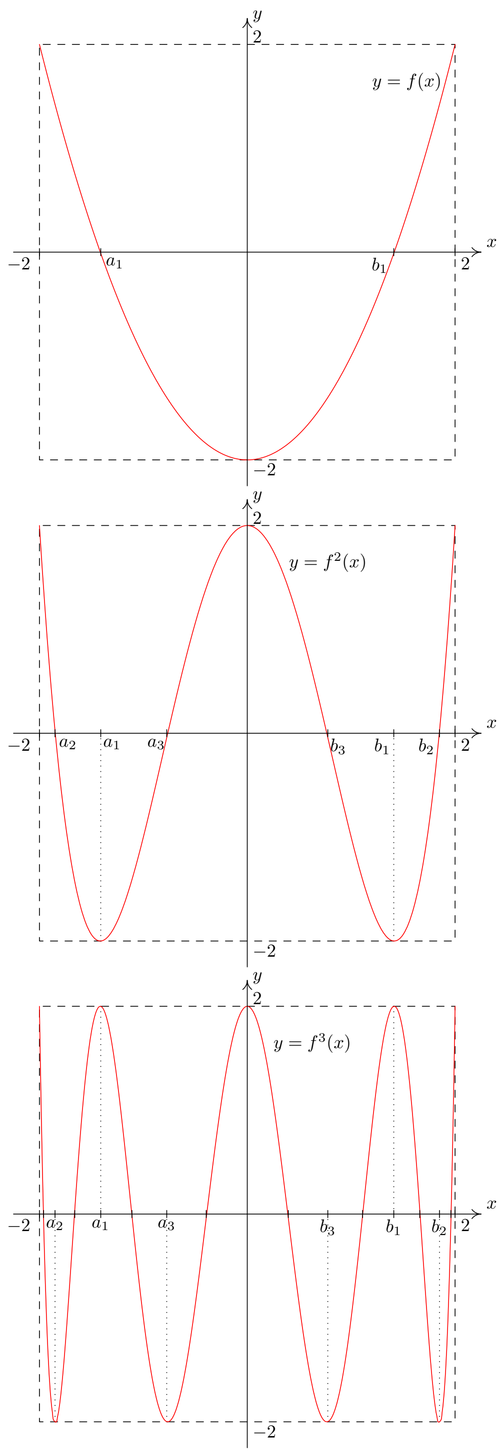

Figure 16.7: The graphs of $f(x) = x^2 - 2$, $x \in [-2, 2]$ and its first two iterations $f^2(x)$ and $f^3(x)$.

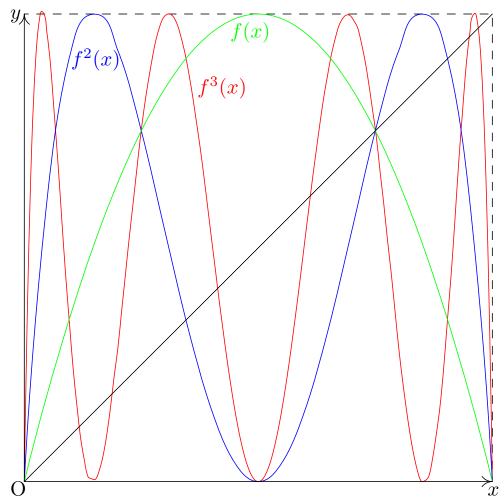

Figure 16.8: The function $f(x) = 4x(1-x)$, $x \in [0,1]$ and its first two iterations $f^2(x)$ and $f^3(x)$.

Solution. The point $x_0 = 0$ is fixed point of $P(x) = 0$. If x_k is a fixed point, then the defining equation requires

$$P(x_k^2 + 1) = x_k^2 + 1 ,$$

that is, $x_{k+1} = x_k^2 + 1$ is a fixed point too. Therefore, the infinite sequence

$$0, \ 1, \ 2, \ 5, \ 26, \ \ldots$$

is a sequence of fixed points for any polynomial sought. In other words, the polynomial $Q(x) = P(x) - x$ vanishes for an infinite number of values and, hence, must vanish identically. Therefore, the only polynomial which satisfies the given conditions is

$$P(x) = x . \qquad \square$$

16.7 A Taste of Chaos

I tried to resist from writing this section as it is slightly more distant from the topic of mathematical competitions and the goals of this book. However, it is tightly related to the topic of iterations and eventually I felt like adding it. But I stay within what has been discussed in the previous sections and I only tease you. Hopefully this brief encounter with chaos will whet your appetite and make you seek further information elsewhere[7].

The logistic function

$$f_\lambda(x) = \lambda \, x \, (1 - x) , \quad x \in [0,1] ,$$

[7] There are many great books written at various levels. For beginners I recommend [40] for a non-quantitative introduction, [35] for a more quantitative, entertaining and illuminating introduction of a broad coverage of related topics, and [41] for a pedagogical course-like presentation of the subject.

is the standard example[8]. The properties of this function depend on the value of the constant λ; so, we have used an additional index λ to differentiate functions with different values of the constant. If $\lambda \in [0,1]$, then $f : [0,1] \to [0,1]$. We will assume that this is the case.

In Figure 16.9, I have plotted the logistic function for the values $\lambda = 0.5$, 1, 1.5, 2, 2.5, 3, 3.5, 4 of the constant (top picture), its first iterate $f_\lambda^2(x)$ for the same values of the constant (middle picture) and its second iterate $f^3(x)$ for the same values of the constant (bottom picture). From the plots we notice the following feature: the iterates develop multiple 'humps' that are less or more pronounced depending on the value of the constant λ. As a result, (a) the number of fixed points at each iteration and (b) the slope of the tangent at the various fixed points (which represents the 'character' of the fixed points) are strongly dependent on the constant λ. In the remaining section, I will attempt to explain the precise dependence of the fixed points on the constant λ. To proceed, the reader should have, at least, fully understood Example 16.1.

Fixed Points of Order 1

We start with the fixed points of $f_\lambda(x)$. These can be found easily:

$$f_\lambda(x_*) = x_* \Rightarrow x_* (\lambda - 1 - \lambda x_*) = 0 \Rightarrow x_*^{(1)} = 0 \quad \text{or} \quad x_*^{(2)} = 1 - \frac{1}{\lambda} .$$

We notice that for $\lambda < 1$ the second fixed point lies outside the domain but for $\lambda > 1$ it is an actual fixed point. For $\lambda_c^{(0)} = 1$ the two fixed points coincide. We thus consider this value of λ as a **critical value** that separates two different behaviors. We say that at $\lambda_c^{(0)} = 1$ the fixed point $x_* = 0$ **bifurcates** (splits) into two fixed points $x_* = 0, 1 - \frac{1}{\lambda}$.

The derivative of $f_\lambda(x)$ is

$$f_\lambda'(x) = \lambda(1 - 2x) ,$$

and

$$f_\lambda'(0) = \lambda , \quad f_\lambda'\left(1 - \frac{1}{\lambda}\right) = -\lambda + 2 .$$

Using the Corollary of page 226, the point $x_*^{(1)}$ is attractive for $\lambda < 1$ and repulsive for $\lambda > 1$. The behavior for the critical value $\lambda_c^{(0)} = 1$ must be inferred by other means: Since the splinter of any point converges to $x_* = 0$, it is still an attractive point. The point $x_*^{(2)}$ is attractive for $\lambda < 3$ and repulsive for $\lambda > 3$. The value $\lambda_c^{(1)} = 3$ is thus another critical value. The behavior of $x_*^{(2)}$ for $\lambda_c^{(1)}$ must be inferred by other means.

[8]The first pedagogical presentation of the logistic function (and other related functions) was given in an influential article by Robert May in [52]. May's presentation uses biology's population growth perspective and summarizes all results without proofs. It is a concise article that the reader should find approachable and useful. In his article, at the end of his opening section, May writes: "The review ends with an evangelical plea for the introduction of these difference equations into elementary mathematics courses, so that students' intuition may be enriched by seeing the wild things that simple nonlinear equations can do." His plea has been more successful. The logistic function can be found even in the first semester introductory calculus courses and, of course, in any book more or less related to the topic.

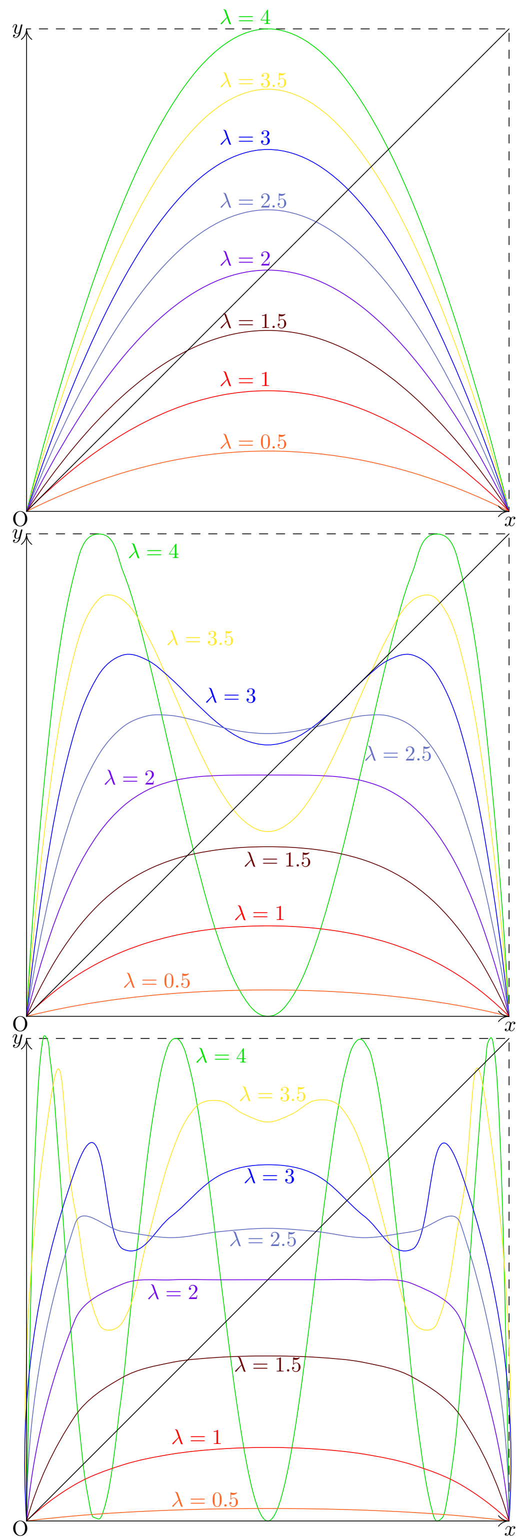

Figure 16.9: The graph of the logistic function $f_\lambda(x)$ and its first two iterations, $f_\lambda^2(x)$, $f_\lambda^3(x)$ for various values of the constant λ. The diagonal line is the graph of the function $y = x$.

2-Cycles

Consider the points a and b such that

$$f(a) = b\,, \quad f(b) = a\,.$$

Therefore, they must be solutions of the equation $f^2(x) = x$,

$$\lambda^2 x(1-x)\,[1-\lambda x(1-x)] = x\,,$$

and not be fixed points of f. For $\lambda \leq 1$, we cannot have such cycles due to Theorem 16.5. Therefore, we must have $\lambda > 1$. Since the two fixed points of order 1 must necessarily be solutions to the previous equation, it can be written in the form

$$x\left[x-\left(1-\frac{1}{\lambda}\right)\right](Ax^2+Bx+C) = 0\,,$$

with the quadratic factor the one determining possible new solutions. Comparing the polynomials term by term, we conclude that the equation which determines the possible 2-cycles is

$$\lambda^2 x^2 - \lambda(\lambda+1)x + (\lambda+1) = 0\,,$$

with discriminant

$$D = \lambda^2(\lambda+1)(\lambda-3)\,.$$

This discriminant is positive for $\lambda > 3$ (2-cycles exist) and negative for $\lambda < 3$ (2-cycles do not exist). When the 2-cycles exist, their points are

$$a, b = \frac{\lambda+1\pm\sqrt{\lambda^2-2\lambda-3}}{2\lambda}\,.$$

The value $\lambda_c^{(1)} = 3$ is the critical value encountered previously. For it, $D = 0$ and we have a double root — a single fixed point, not a 2-cycle:

$$x_* = \frac{2}{3} = 1-\frac{1}{\lambda_c^{(1)}}\,.$$

That is, the point $x_*^{(2)}$. For $\lambda > 3$ the fixed point $x_*^{(2)}$ bifurcates in two points that are not fixed points of f but they constitute a 2-cycle. The discriminant for the 2-cycle is

$$D_{2-cycle} = f'_\lambda(a)\,f'_\lambda(b) = \lambda^2\,(1-2a)\,(1-2b) = -\lambda^2+2\,\lambda+4\,,$$

and takes the value $D_{2-cycle} = -1$ for $\lambda_c^{(2)} = 1+\sqrt{6}$. If $\lambda < 1+\sqrt{6}$ the 2-cycle is attractive; if $\lambda > 1+\sqrt{6}$ the 2-cycle is repulsive. Therefore, we discovered a new critical point $\lambda_c^{(2)} = 1+\sqrt{6} \simeq 3.4495$.

2^n-Cycles

Continuing in the same fashion as above, we can search for 4-cycles, 8-cycles, and so on. Since the calculations become exceedingly harder however, analytical results are harder to obtain, so one relies on the help of computers. In this way, we can discover that for $\lambda > 1+\sqrt{6}$, a 4-cycle emerges. This 4-cycle is attractive for $\lambda < \lambda_c^{(3)} \simeq 3.5441$ and is repulsive for $\lambda > \lambda_c^{(3)}$.

For $\lambda > \lambda_c^{(3)}$, an 8-cycle emerges. This 8-cycle is attractive for $\lambda < \lambda_c^{(4)} \simeq 3.5644$ and is repluslive for $\lambda > \lambda_c^{(4)}$.

For $\lambda > \lambda_c^{(4)}$, a 16-cycle emerges. This 16-cycle is attractive for $\lambda < \lambda_c^{(5)} \simeq 3.5688$ and is repluslive for $\lambda > \lambda_c^{(5)}$.

This procedure of **period doubling** continues indefinitely and creates the increasing sequence

$$\lambda_c^{(0)} < \lambda_c^{(1)} < \lambda_c^{(2)} < \lambda_c^{(3)} < \cdots \le 4\,.$$

Since the sequence is bounded, it converges to a limit

$$\lim_{n\to\infty} \lambda_c^{(n)} \;=\; \lambda_c \simeq 3.5699\,.$$

Bifurcation Diagram

To prepare the reader for a more complicated plot, we have plotted the bifurcations of the logistic function up to 4-cycles in Figure 16.10. More precisely, the vertical axis in this figure gives the value of the fixed point (of any order) versus the value of the constant λ (horizontal axis). A change in color indicates a bifurcation, that is the appearance of a new cycle with double period. The dashed lines indicate repulsive points and cycles.

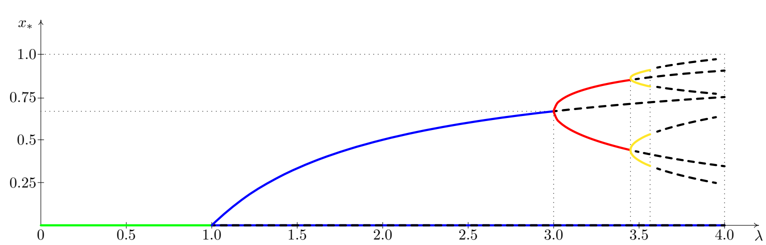

Figure 16.10: Partial bifurcation diagram for logistic function $f_\lambda(x)$. Often the dashed lines are omitted. In such a case, the diagram is called an **orbit diagram**.

The bifurcations for the cycles come faster and faster, that is the **bifurcation distances** $\Delta\lambda_n = \lambda_c^{(n+1)} - \lambda_c^{(n)}$ decrease:

$$\Delta\lambda_1 > \Delta\lambda_2 > \Delta\lambda_3 > \ldots\,,$$

Even more, Feigenbaum observed, the decrease is almost similar to that of a geometric progression: $\Delta\lambda_n \simeq \delta\,\Delta\lambda_{n+1}$. This approximate equation can become exact,

$$\lim_{n\to\infty} \frac{\Delta\lambda_n}{\Delta\lambda_{n+1}} = \delta \,,$$

where $\delta \simeq 4.669$.The number δ is thus known as **Feigenbaum's constant**.

Similarly, we can define a decreasing sequence of **bifurcation widths** Δw_n,

$$\Delta w_1 > \Delta w_2 > \Delta w_3 > \dots \,,$$

that measures the widths of the lower forks at the bifurcation points. Then it can be proved that this sequence is also similar to that of a geometric progression:

$$\lim_{n\to\infty} \frac{\Delta w_n}{\Delta w_{n+1}} = \alpha \,,$$

where $\alpha \simeq 2.5029$.

3-Cycles

Now, let's return to the 3-cycles. Consider the points a, b and c such that

$$f(a) = b \,, \quad f(b) = c \,, \quad f(c) = a \,.$$

Therefore, they must be solutions of the equation $f^3(x) = x$,

$$\lambda^3 x(1-x)\,[1-\lambda x(1-x)]\left\{1-\lambda^2 x(1-x)\,[1-\lambda x(1-x)]\right\} = x \,.$$

Since the fixed points of f will necessary be roots of this equation, it should factorize in the form

$$x\left(x-1+\frac{1}{\lambda}\right)P(x) = 0 \,,$$

where $P(x)$ is a polynomial of degree 6.

For some values λ the polynomial $P(x)$ has no real roots; for some other values it has six real roots. In the critical case, $\lambda = \lambda_0$ that separates these two domains, the polynomial should have three real double roots. (See Figure 16.11.) Therefore we write

$$P(x) = [(x-\alpha)(x-\beta)(x-\gamma)]^2 = (x^3 - Ax^2 + Bx - C)^2 \,,$$

where

$$A = \alpha+\beta+\gamma \,, \quad B = \alpha\beta+\beta\gamma+\gamma\alpha \,, \quad C = \alpha\beta\gamma \,.$$

Comparing the two forms of the polynomial, we can find A, B, C as functions of λ,

$$\begin{aligned} A &= \frac{3\lambda+1}{2\lambda} \,, \\ B &= \frac{(3\lambda+1)(\lambda+3)}{8\lambda^2} \,, \\ C &= \frac{-\lambda^3+7\lambda^2+5\lambda+5}{16\lambda^3} \,. \end{aligned}$$

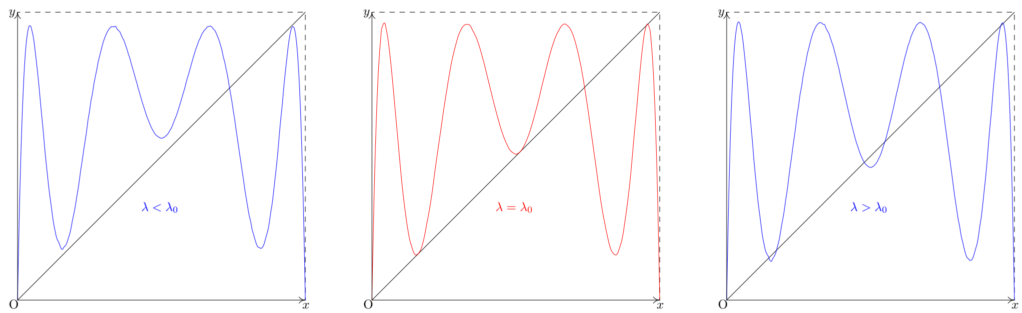

Figure 16.11: Variation of the parameter λ to find the critical value λ_0 at which 3-cycles start to appear. At the critical value f_λ^3 is tangent to the diagonal at three points. Below the critical value the curve 'pulls away' from these points, while above the critical value each contact point bifurcates to create two new points.

In addition we find the three conditions

$$g_1(\lambda)\,(\lambda^2 - 2\lambda - 7) \;=\; g_2(\lambda)\,(\lambda^2 - 2\lambda - 7) \;=\; g_3(\lambda)\,(\lambda^2 - 2\lambda - 7) \;=\; 0\,,$$

where g_1, g_2, g_3 are some functions for which the exact form is not important. Obviously, the last conditions admit the common solution $\lambda^2 - 2\lambda - 7 \;=\; 0$, or $\lambda \;=\; 1 \pm \sqrt{8}$. From the two values of λ thus obtained, only the value $\lambda_0 = 1 + \sqrt{8} \simeq 3.8284$ is acceptable. Given this value of λ, one can then find easily A, B, C and, from them after solving a cubic equation, the roots α, β, γ.

From Šarkovskii's theorem we know that, if there are 3-cycles, there will be cycles of any period. One can proceed to find them by the same techniques used for the 2^n-cycles.

Chaos

The word *chaos* is derived from the Greek word χάος meaning *absolute disorder*. It was introduced in the influential paper of Li and Yorke [51] entitled *Period Three Implies Chaos*. It really means existence of non-periodic orbits for which Li and Yorke established a criterion. The term became very popular and today *chaos theory*[9] is a term used even by layman.

I will avoid precise definitions in this subsection. Instead I will use a loose, heuristic discussion. So, given the function $f : I \to I$, consider an uncountable set $S \subseteq I$ of points of f which satisfies the following two conditions:

- The splinter of a point belonging to S does not approach asymptotically any periodic orbit.
- The splinters of two points belonging to S do not approach each other asymptotically but they can come arbitrarily close.

[9] I have always found the term *chaos* very misleading and inappropriate since it implies lack of control and knowledge. Yet, the subject is about deterministic equations.

Such a set S is called a **scrambled set**.

If a function f has such a **scrambled set**, it is called **Li-Yorke chaotic**. Then, Li and Yorke showed that the condition "f has a 3-cycle" is a sufficient one for f to be Li-Yorke chaotic. After Li and Yorke's work, their result was extended and generalized. In particular, one can actually show that it suffices for f to have a $(2m+1)2^n$-cycle for some m, n, $m \geq 1$ to be Li-Yorke chaotic. Even more, a function with 2^n-cycles of any order but no $(2m+1)2^n$-cycles with $m \neq 0$ can also be Li-Yorke chaotic.

It turns out that the Li-Yorke chaos provides only a limited aspect of the study of chaos in a system. The reason is that, if all scrambled sets are of measure zero, their probability of choosing an initial point in these sets is zero. Therefore, although the Li-Yorke chaos may be present, *it is unobservable* in such a scenario. Therefore, other criteria and definitions of chaos — in particular, criteria for *observable* chaos — which are not necessarily equivalent to Li-Yorke chaos have been proposed. Such a criterion is, for example, the **Lyapunov exponent** that measures sensitivity on the initial conditions. Given two nearby points x and y separated by distance d_0, under n iterations they will find themselves separated by distance d_n. If we write

$$d_n \;=\; d_0\, e^{nL_n} \;,$$

the number L_n is the Lyapunov exponent for the n-th iterate. A positive L implies that the two points will separate exponentially fast and it is considered a footprint of chaos.

Exercise 16.1. *Assuming that the limit* $\lim\limits_{n\to\infty} L_n$ *exists and it is* L*, show that*

$$L \;=\; \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln|f'(x_k)| \,,$$

and that for attractive cycles $L < 0$. □

Orbit Diagram, Self-Similarity, & Universality

Figure 16.12 shows the orbit diagram (that is, only the attracting points) of the logistic function. I will make some comments to reveal the mixture of order and chaos present in this diagram.

The most evident feature is the 'unshaded' regions known as the **periodic windows**. There is a countable but infinite number of such windows and they have a finite width. For numerical analysis, it is know that the sum of the widths of all windows is about 10% of $4 - \lambda_c$.

It looks as inside the periodic windows the splinter of any point x approaches a cycle (that, is we have 'order') and outside these windows the splinter behaves chaotically, taking any possible value in $[0, 1]$. However, the situation is more complicated (and exciting). Imagine that we magnify a part of the orbit diagram, say the 'upper branch' from

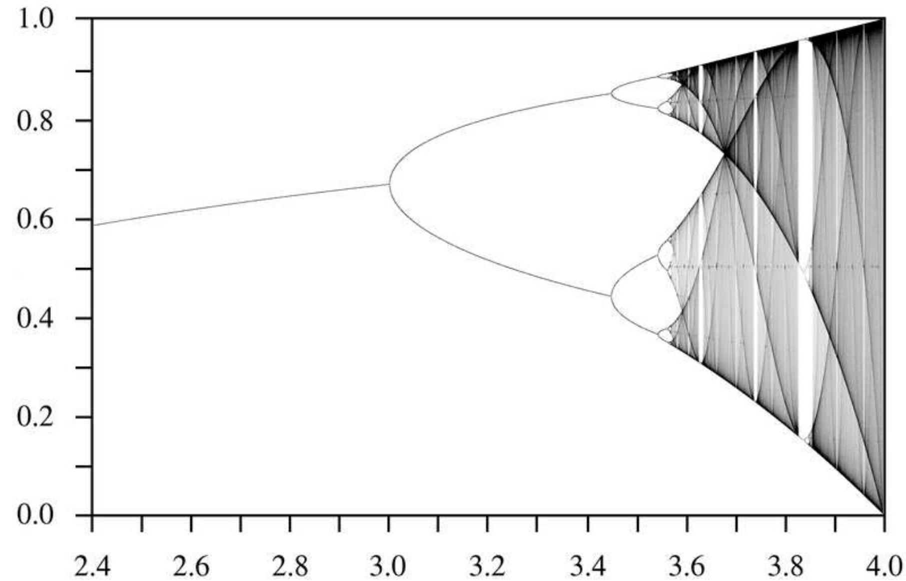

Figure 16.12: The orbit diagram of the logistic function $f_\lambda(x)$.

$\lambda = 3.2$ to $\lambda = 3.65$. Then the emerging picture would be indistinguishable of the original orbit diagram of Figure 16.12. As such, we can magnify again to find another identical part, and so on ad infinitum. We have uncovered an amazing self-similar pattern.

To be more concrete, let's, for example, look at the window that starts at $\lambda_0 \simeq 3.8284$. This is the *period-three window* which is unique with this period and it has the largest width among windows. Given a λ in the window and *almost* any point x, the splinter of x will approach asymptotically the 3-cycle. The word 'almost' is used here since there are points x for which their splinters do not approach asymptotically the 3-cycle. This can be understood as follows. Inside the period-three window there is a copy of the original orbit diagram: In the same way that period doubling led to the stable periodic orbits of period 2^n in the original orbit diagram, period doubling inside the period-three window leads to stable periodic orbits of periods $3 \cdot 2^n$. Then, inside the period-three window there is a period-nine window, which in the magnification will appear as a period-three window. And so on. The non-periodic orbits that are found inside the period-three window provide a 'small' Li-Yorke chaos in this window but they have measure zero and thus they cannot be observed.

Feigenbaum discovered that the qualitative features of the function (that is everything we have concluded except the numerical values) are *identical* for all functions (not only the logistic function) that are concave with a single maximum (known as **unimodal** functions).

Examples of such functions are

$$\begin{aligned} f_a(x) &= a \sin(\pi x) \,, \\ f_b(x) &= b - x^2 \,. \end{aligned}$$

Although numerical values are different from function to function (e.g. the exact bifurcation points $\lambda_c^{(0)}, \lambda_c^{(1)}, \lambda_c^{(2)}, \ldots$ are different), the constants α and δ are however the same for all such functions. This means, that Feigenbaum's constant plays the same role in the theory of unimodal functions as the role of Archimedes' constant in the theory of the circle.

Based on the previous observations and a mathematical framework borrowed from physics[10], Feigenbaum built a beautiful mathematical formalism to explain the appearance of the constant δ and α. Unfortunately any attempt to explain more on the topic would take us very deeply into mathematics and physics. Hopefully, the reader has appreciated the rich behavior of even simple functions and the deep connections among various areas of modern mathematics and physics.

Problem 16.10. *The functional equation*

$$\alpha f^2(x) = f(\alpha x) \,, \quad \alpha \neq 0$$

is called the ***Feigenbaum equation****. It is encountered in Feigenbaum's theory for the explanation of the universal properties of unimodal functions. Find as many continuous solutions $f : \mathbb{R} \to \mathbb{R}$ as you can. (For such solutions see [53].)*

A Quiz

First attempt to solve the following problem:

Problem 16.11 (USA 1974)**.** *Let a, b, c denote three distinct integers and $P(x)$ denote a polynomial with integer coefficients. Show that it is impossible that $P(a) = b$, $P(b) = c$, and $P(c) = a$.*

That is, $P(x)$ does not have a 3-cycle. In fact, you can try to explore the validity of the statement for any k-cycle.

Then, notice that the function

$$f(x) = 4x(1-x) = -4x^2 + 4x + 0 \,,$$

is a polynomial with integer coefficients. Moreover, we have shown that it has cycles of any order. Was the problem incorrectly given before the publication of Li and Yorke's paper in 1975 and knowledge of Šarkovskii's theorem? Explain.

[10] This framework is known as **renormalization** and it plays a fundamental role in the description and explanation of critical phenomena in condensed matter physics and the unification of forces in high energy physics.

Chapter 17

Solving by Invariants & Linearization

In this chapter we shall consider functional equations of the form

$$f(h(x, f(x))) = H(x, f(x)) , \tag{17.1}$$

where $f(x)$ is the unknown function and $h(u, v), H(u, v)$ are given functions of two variables. This equation includes as special cases all equations we have discussed in the last chapter. In particular, the Abel (16.3), Schröder (16.4), and Böttcher (16.5) equations are of the form

$$f(h(x)) = H(f(x)) \tag{17.2}$$

which is known as **conjugacy equation**. When there is an injective function $f(x)$ such that the above equation is true, we call the functions h, H **conjugate**.

17.1 Constructing Solutions

Characteristic Function

Consider now the graph $\Gamma_f \subseteq \mathbb{R}^2$ of the function $f : A \to A$ that solves equation (17.1):

$$\Gamma_f = \{(x, f(x)) , x \in A\} .$$

If $x \in A$, then $h(x, f(x)) \in A$ too. So if the point $(x, f(x))$ belongs to the graph, so does the point $(h(x, f(x)), f(h(x, f(x))))$. Using (17.1), we can write the latter point as $(h(x, f(x)), H(x, f(x)))$.

Motivated by the above, let's define a function $\chi : \mathbb{R}^2 \to \mathbb{R}^2$ such that $(x, y) \mapsto (x', y')$ such that

$$\begin{aligned} x \mapsto x' &= h(x, y) , \\ y \mapsto y' &= H(x, y) . \end{aligned}$$

Thus defined, this function maps Γ_f to itself and it is known as the **characteristic function** of the equation (17.1). For the conjugacy equation (17.2), the characteristic function is

$$\begin{aligned} x \mapsto x' &= h(x) , \\ y \mapsto y' &= H(y) . \end{aligned}$$

Invariants

Definition 17.1. Given the function $\chi : \mathbb{R}^2 \to \mathbb{R}^2$ and a function $\Phi(u, v)$ of two variables, the function Φ is called an **invariant** of χ if $\Phi(u, v) = \Phi(\chi(u, v))$ for all $(u, v) \in \mathbb{R}^2$.

Definition 17.2. Given the function $\chi : \mathbb{R}^2 \to \mathbb{R}^2$ and a subset A of the plane, A is called an **invariant set** with respect to χ if $\chi(A) \subseteq A$.

Some immediate comments are:

1. If Φ is an invariant, then the locus
$$A_c \;=\; \{(u, v) \in \mathbb{R}^2 \mid \Phi(u, v) = c\}\;,$$
where c is a number, is an invariant set.
2. If $\Phi_1(u, v)$ and $\Phi_2(u, v)$ are two invariants of the same function χ, then $F(\Phi_1(u, v), \Phi_2(u, v))$ is also an invariant of χ for any function $F(u, v)$.
3. If A_1 and A_2 are two invariant sets of the same function χ, then their intersection $A_1 \cap A_2$ is an invariant of χ too.

There exists a theory for the construction of invariants but it is beyond the goal of this book. We will rely on simple constructions only. For example, if the function χ is **cyclic**, that is $\chi^k = \mathrm{id}$ for some k, we can can construct invariants as follows: Select any function $\phi(u, v)$ and define
$$\phi_i(u, v) \;=\; \phi(\chi^i(u, v))\;, \quad i = 1, 2, \ldots, k-1\;.$$
Then
$$\Phi(u, v) \;=\; \phi(u, v)\phi_1(u, v)\phi_2(u, v)\ldots\phi_{k-1}(u, v)$$
is an invariant of χ. For other simple cases, we can construct invariants by simple inspection.

Finding Solutions

Given a functional equation (17.1), we may write down its characteristic function and construct an invariant $\Phi(x, y)$. Then the set of points $\Phi(x, y) = c$ may give a solution to the functional equation which can be found by solving for $y = f(x)$. We say 'may' since an equation may have an invariant but there may not be a solution.

The above statement can be understood better with an example. Let's construct at least one solution of the equation
$$f(f(x) - 2x) \;=\; 2f(x) - 3x\;.$$
The corresponding characteristic function χ is
$$\begin{aligned} x \mapsto x' &= y - 2x\;,\\ y \mapsto y' &= 2y - 3x\;. \end{aligned}$$

It's easy to check that this function is cyclic: $\chi^2 = \mathrm{id}$. So, let's pick the function $\phi(x, y) = x$ from which we can construct the invariant

$$\Phi(x, y) = x(y - 2x) .$$

Equating $\Phi(x, y)$ to a number c, results to

$$x(y - 2x) = c \Rightarrow y = \frac{c}{x} + 2x .$$

By substitution into the original equation, we can easily verify that the function

$$f(x) = \frac{c}{x} + 2x$$

is indeed a solution.

17.2 Linear Equations

A **linear functional equation** is one of the form

$$f(g(x)) = h(x)\, f(x) + F(x) , \tag{17.3}$$

where $f : I \to I$ is the unknown function and $g, h, F : I \to I$ are known functions with $g(I) \subseteq I$. When $F(x) \not\equiv 0$, the equation is called **non-homogeneous**; when $F(x) \equiv 0$,

$$f(g(x)) = h(x)\, f(x) , \tag{17.4}$$

and it is called **homogeneous**.

The Schröder equation (16.4) is a homogeneous linear equation with $h(x) = \lambda$; the Abel equation (16.3) is a non-homogeneous linear equation with $h(x) = 1$ and $F(x) = a$.

A linear homogeneous equation has the important property that given any two solutions, their sum is also a solution. The following theorem provides an alternative way to construct additional solutions to (17.4), given that one solution is already known.

Theorem 17.1. *Let $f_0(x)$ be a solution of (17.4). If $\Phi(x)$ is an invariant of its characteristic function that depends only on the first variable, then*

$$f(x) = \Omega(\Phi(x))\, f_0(x) ,$$

where Ω is an arbitrary function, is also a solution of (17.4).

Proof. Since $\Phi(x)$ is an invariant, $\Phi(g(x)) = \Phi(x)$ and therefore $\Omega(\Phi(g(x))) = \Omega(\Phi(x))$. So

$$\begin{aligned} f(g(x)) &= \Omega(\Phi(g(x)))\, f_0(g(x)) \\ &= \Omega(\Phi(x))\, h(x)\, f_0(x) \\ &= h(x)\, f(x) , \end{aligned}$$

as required. ◻

As is true with linear difference equations, the solutions of the non-homogeneous equation (17.3) are obtained from those of the homogeneous one by the addition of a particular solution:

Theorem 17.2. *Let $f_p(x)$ be a particular solution of the non-homogeneous equation (17.3). Then*

(a) if $f_0(x)$ is a solution of the homogeneous equation (17.4), the function $f(x) = f_p(x) + f_0(x)$ is a solution of the non-homogeneous equation (17.3).

(b) every solution of the non-homogeneous equation (17.3) can by obtained from $f_0(x)$ in the way stated in part (a).

Proof. (a) It is easy to verify this statement by direct substitution in the right hand side of (17.3):

$$h(x)\,(f_p(x) + f_0(x)) + F(x) \;=\; [h(x)\,f_p(x) + F(x)] + [h(x)\,f_0(x)] \;=\; f_p(g(x)) + f_0(g(x))\,.$$

(b) Let $f_p(x)$ be some solution of the non-homogeneous equation (17.3) and $f(x)$ any other solution. Then

$$\begin{aligned} f(g(x)) &= h(x)\,f(x) + F(x)\,, \\ f_p(g(x)) &= h(x)\,f_p(x) + F(x)\,. \end{aligned}$$

By subtracting the two equations we find

$$(f - f_p)(g(x)) \;=\; h(x)\,(f - f_p)(x)\,;$$

that is, the difference $f - f_p$ must be a solution to the homogeneous equation (17.4). □

17.3 The Abel and Schröder Equations

The Abel and Schröder equations play a central role in the theory of functional equations without 'parameters'. Their importance relies on the fact that they consist the simplest linear equations whose solutions are used in the construction of solutions of more complicated equations. This will become obvious in the remaining chapter.

Usually, we choose $a = 1$ in Abel's equation:

$$f(g(x)) \;=\; f(x) + 1\,. \tag{17.5}$$

Assume that f, g are defined on some subset A of $\mathbb{R}$.

Theorem 17.3. *Abel's equation has a solution iff the function $g(x)$ has neither fixed points nor cycles. In such a case, it has infinitely many solutions.*

Proof. Setting $g(x)$ in place of x we find

$$f(g^2(x)) \;=\; f(g(x)) + 1 \;=\; f(x) + 2\,.$$

Repeating this procedure

$$f(g^n(x)) = f(x) + n\,, \quad n \in \mathbb{N}^*\,. \tag{17.6}$$

If g has a cycle of order n ($n = 1$ being the case of a fixed point) for the point x_0, then $f(g^n(x_0)) = f(x_0)$ and we thus arrive at the contradiction $0 = n$.

Let's consider now the orbits generated by the iteration of g. (See section 16.2.) For each orbit O_a we select a representative element $a \in O_a$. Also, for simplicity, let $b = f(a)$. For any point $a' \in O_a$, there exist some $n, n' \in \mathbb{N}^*$ such that $g^n(a) = g^{n'}(a')$. Then using equation (17.6):

$$f(g^{n'}(a')) = f(g^n(a)) \Rightarrow f(a') = b + n - n'\,.$$

That is, the function f is known on all points of the orbit if it is known at one point. From this we conclude that the function f is fully defined when one value b is given for one point a of each of the orbits generated by g. Since there are an infinite number of ways to assign $a \mapsto b$, the theorem is proved. □

Now let's turn our attention to the Schröder equation (16.4). We can notice immediately that for $\lambda > 0$, $\lambda \neq 1$, if we exponentiate Abel's equation (17.5)

$$\lambda^{f(g(x))} = \lambda\, \lambda^{f(x)}\,,$$

any solution $f(x)$ of Abel's equation gives a solution $\lambda^{f(x)}$ for the Schröder equation too. Of course, there may be additional solutions not obtained in this way and, obviously, this construction does not provide any solution for $\lambda < 0$ (if there are such solutions).

17.4 Linearization

Imagine that a (non-linear) function $g(x)$ is given. Consider also that we can find a continuous monotonic solution $\sigma(x)$ of the Schröder equation

$$\sigma(g(x)) = \lambda\, \sigma(x)\,, \quad \lambda \neq -1, 0, 1\,.$$

We can use the above data to simplify equation (17.3). We introduce the change of variables

$$\begin{aligned} y &= \sigma(x)\,, \\ \tilde{f}(y) &= f(x)\,. \end{aligned}$$

Since σ is monotonic, it has an inverse σ^{-1} and $x = \sigma^{-1}(y)$. Also we notice that, if we replace x by $g(x)$ in the second equation $\tilde{f}(\sigma(x)) = f(x)$, we have $\tilde{f}(\sigma(g(x))) = f(g(x)) \Rightarrow \tilde{f}(\lambda\sigma(x)) = f(g(x)) \Rightarrow \tilde{f}(\lambda y) = f(g(x))$. Substituting these results into (17.3), we find:

$$\tilde{f}(\lambda y) = h(\sigma^{-1}(y))\, \tilde{f}(y) + F(\sigma^{-1}(y))\,.$$

We can make this equation more appealing with the definition of the new functions:

$$\tilde{h}(y) \ = \ h(\sigma^{-1}(y)) \ , \quad \tilde{F}(y) \ = \ F(\sigma^{-1}(y)) \ .$$

Then,

$$\tilde{f}(\lambda y) \ = \ \tilde{h}(y) \, \tilde{f}(y) + \tilde{F}(y) \ . \tag{17.7}$$

Equation (17.7) has the same form as (17.3) but a simpler function $\tilde{g}(y) = \lambda y$ appears in the left hand side. It may thus be simpler to solve than the original equation.

Similarly, if a monotonic solution $\sigma(x)$ of Abel's equation is used, we can make a similar change of variables in (17.3) to arrive at a simpler equation:

$$\tilde{f}(y+1) \ = \ \tilde{h}(y) \, \tilde{f}(y) + \tilde{F}(y) \ , \tag{17.8}$$

with $\tilde{g}(y) = y + 1$.

Example 17.1. Consider the equation

$$f((x-a)^2 + a) \ = \ b \, f(x) \ ,$$

with a, b some constants and $b > 0$.

Instead of x, it is more convenient to use $X = x - a$:

$$f(X^2 + a) \ = \ b \, f(X + a) \ .$$

The function $g(X) = X^2 + a$ is quadratic. So let's consider a monotonic solution of the Schröder equation

$$\sigma(X^2) \ = \ 2\sigma(X) \ ,$$

with solution $\sigma(X) = \log X$.

So we introduce the change of variables:

$$\begin{aligned} y \ &= \ \ln X \ , \\ \tilde{f}(y) \ &= \ f(X + a) \ . \end{aligned}$$

Since $X = e^y$, we see that $\tilde{f}(y) = f(e^y + a)$ and the original equation transforms to

$$\tilde{f}(2y) \ = \ b \, \tilde{f}(y) \ .$$

This equation has an obvious solution

$$\tilde{f}(y) \ = \ y^{\log_2 b} \ .$$

Returning to the original notation

$$f(x) \ = \ (\ln(x-a))^{\log_2 b} \ . \qquad \square$$

It is tempting to state that the above method of reduction to a simpler equation, known as *linearization*, can be applied to the more general functional equation

$$F(x, f(x), f(g(x))) = 0 ,$$

where F, g are known functions and f is the unknown one. Assume that one can find bijective functions σ and τ such that

$$\begin{aligned} \sigma(g(x)) &= \tilde{g}(\sigma(x)) , \\ \tau(F(x, y, z)) &= \tilde{F}(x, \tau(y), \tau(z)) . \end{aligned}$$

The first equation is the conjugacy equation (17.2) and it replaces the function g by its conjugate $\tilde{g}$ which is, hopefully, a simpler function. The second equation is a generalization of the conjugacy equation to a function of more variables — notice the way the arguments are treated based on their interpretation. Upon the change of variables

$$\begin{aligned} x' &= \sigma(x) , \\ \tilde{f} &= \tau \circ f \circ \sigma^{-1} , \end{aligned}$$

the original functional equation transforms to

$$\tilde{F}(\sigma^{-1}(x'), \tilde{f}(x'), \tilde{f}(\tilde{g}(x'))) = 0 ,$$

which might be easier to solve. This equation is a **linearization** of the original equation if $\tilde{F}, \tilde{g}$ are linear functions.

17.5 Solved Problems

The two problem that follow are taken from [39] and they complete, more or less, the necessary ideas the reader should know about the topic at this introductory level.

Problem 17.1. *Find at least one solution of the equation*

$$f(x^2 - 2f(x)) = f(x)^2 .$$

Solution. The characteristic function χ of the given equation is

$$\begin{aligned} x \mapsto x' &= x^2 - 2y , \\ y \mapsto y' &= y^2 . \end{aligned}$$

We write down the function

$$\Phi(x, y) = \frac{x^2}{y} .$$

It is almost apparent that this function is not invariant under χ but it transforms as

$$\Phi(\chi(x,y)) = \frac{(x^2-2y)^2}{y^2} = (\Phi(x,y)-2)^2 .$$

On the other hand the function

$$g(x) = (x-2)^2$$

has two fixed points

$$g(x) = x \Rightarrow x^*_\pm = 1, 4 ,$$

that can allow us to define two invariant sets:

$$\begin{aligned} A_+ &= \{(x,y) \mid \Phi(x,y) = 4\} , \\ A_- &= \{(x,y) \mid \Phi(x,y) = 1\} . \end{aligned}$$

The set A_+ is the graph of the function

$$\Phi(x,y) = 4 \Rightarrow y = \frac{x^2}{4} ,$$

while the set A_- is the graph of the function

$$\Phi(x,y) = 1 \Rightarrow y = x^2 .$$

It can be verified by direct substitution that both functions are solutions of the given functional equation.

The previous construction begs an interesting question: if a function $\Phi(x,y)$ is not invariant but has cycles, could we use them to find solutions? In particular, $g(x) = (x-2)^2$ has a 2-cycle. Can it lead to a solution of the given equation? As we demonstrate below, the answer is affirmative.

To find the 2-cycle, we search for fixed points of $g^2(x) = [(x-2)^2-2]^2$:

$$g^2(x) = x .$$

Let a_* be a solution of this equation. (Show that there is such a point.) Then we define $b_* = g(a_*) = (a_*-2)^2$. The two numbers a_* and b_* are the distinct points of a 2-cycle. Now consider the set

$$A = \{(x,y) \mid (\Phi(x,y)-a_*)(\Phi(x,y)-b_*) = 0\} .$$

This defines an invariant set. To see this, if a point belongs to A, then $\Phi(x,y)$ takes one of the two values of the 2-cycle, say a_*. So $\Phi(x,y) = a_*$ and

$$\Phi(\chi(x,y)) - b_* = (\Phi(x,y)-2)^2 - b_* = (a_*-2)^2 - b_* = 0 ,$$

that is, $\Phi(\chi(x, y))$ also belongs to A. We thus find the solution

$$f(x) \;=\; \begin{cases} \frac{x^2}{a_*}\,, & \text{if } x \geq 0\,, \\ \\ \frac{x^2}{b_*}\,, & \text{if } x < 0\,, \end{cases}$$

or the solution

$$f(x) \;=\; \begin{cases} \frac{x^2}{b_*}\,, & \text{if } x \geq 0\,, \\ \\ \frac{x^2}{a_*}\,, & \text{if } x < 0\,. \end{cases} \qquad \square$$

Problem 17.2. *Find a solution $f : (1, +\infty) \to \mathbb{R}_+^*$ of the functional equation*

$$f(x^a) \;=\; \lambda\, x^b\, f(x)\,,$$

where $a, \lambda \in \mathbb{R}_+^*$, $b \in \mathbb{R}$.

Solution. Linearization: Consider the Schröder equation

$$\sigma(x^a) \;=\; a\,\sigma(x)\,,$$

which admits the solution $\sigma(x) = \ln x$. Upon the change of variables

$$\begin{aligned} x' \;&=\; \ln x\,, \\ \tilde{f}(x') \;&=\; f(x)\,, \end{aligned}$$

the original functional equation transforms to

$$\tilde{f}(ax') \;=\; \lambda\, e^{bx'}\, \tilde{f}(x')\,.$$

To solve this new equation we try the method of invariants as shown in the next part of the solution. In particular, a modified version of the method of invariants is presented that can be often useful.

Solution by invariants: We write the function $\tilde{f}(x')$ as the ratio of two other functions:

$$\tilde{f}(x') \;=\; \frac{A(x')}{B(x')}\,.$$

Then

$$\frac{A(ax')}{B(ax')} \;=\; \frac{\lambda}{e^{-bx'}}\,\frac{A(x')}{B(x')}\,.$$

Finding a solution $\tilde{f}(x')$ can now be reduced to finding solutions to two simpler equations:

$$A(ax') = \lambda A(x'), \tag{17.9}$$
$$B(ax') = e^{-bx'} B(x'). \tag{17.10}$$

We solve each equation using the method of invariants.

Equation for A: The characteristic function of equation (17.9) is

$$\begin{aligned} x' \mapsto x'' &= a\,x', \\ y \mapsto y' &= \lambda\, y. \end{aligned}$$

Instead of using χ directly, we modify it by applying a suitable function. So, we apply the logarithmic function to find

$$\begin{aligned} \ln x' \mapsto \ln x'' &= \ln a + \ln x', \\ y \mapsto \ln y' &= \ln \lambda + \ln y. \end{aligned}$$

We now see that the functions

$$\begin{aligned} \Phi_1(x', y) &= \frac{\ln x'}{\ln a} - \frac{\ln y}{\ln \lambda}, \\ \Phi_2(x', y) &= \Omega\left(\frac{\ln x'}{\ln a}\right), \end{aligned}$$

where Ω is a periodic function with period 1 are invariants. Setting $\Phi_1(x, y) = 0$ we find the solution

$$A_0(x') = (x')^{\log_a \lambda}.$$

We can combine this solution with the second invariant to construct other solutions of the equation according to Theorem 17.1:

$$A(x') = (x')^{\log_a \lambda}\, \Omega\left(\frac{\ln x'}{\ln a}\right).$$

Equation for B: We will solve equation (17.10) along similar lines. Its characteristic function is

$$\begin{aligned} x' \mapsto x'' &= a\,x', \\ y \mapsto y' &= e^{-bx'}\, y. \end{aligned}$$

This we modify as follows

$$\begin{aligned} x' \mapsto x'' &= a\,x', \\ \ln y \mapsto y' &= -b\,x' + \ln y. \end{aligned}$$

From this, we can see that

$$\Phi(x', y) = \ln y + \frac{b}{a-1}\, x'$$

is an invariant. Setting $\Phi(x', y) = 0$ gives the solution

$$B(x') = e^{\frac{b}{1-a}x'} .$$

A solution to the original equation: Using the functions A and B, we now have solution of the linearized equation:

$$\tilde{f}(x') = (x')^{\log_a \lambda} e^{-\frac{b}{1-a}x'} \Omega(\log_a x') .$$

Returning to the original notation and the initial problem, the function

$$f(x) = (\ln x)^{\log_a \lambda} x^{\frac{b}{a-1}} \Omega(\log_a \ln x) ,$$

where Ω a periodic function with period 1, is a solution of the given functional equation. $\square$

Chapter 18

More on Fixed Points

Although we defined the concept of a fixed point for a function in one variable, the concept is much, much deeper and permeates all topics in mathematics. The concept really belongs to the area of topology[1] — a branch of mathematics that originated with the work of Poincaré dealing with and characterizing the geometric properties of any set.

Fixed point theorems are used as a tool to prove existence theorems in differential, integral, and other equations. They are used in dynamical systems and their applications to game theory, economics, biology, etc.

The reader may look at Shashkin's book [36] for a nice and short review of the use of fixed points. In this chapter, we only demonstrate how fixed points can be used to solve functional equations which contain parameters.

18.1 Solved Problems

Problem 18.1 (IMO 1983)**.** *Find all functions f defined on the set of positive real numbers which take positive real values and satisfy the conditions*

(a) $f(xf(y)) = yf(x)$, for all positive x, y;

(b) $f(x) \to 0$ as $x \to +\infty$.

Solution. The function $f(x) \equiv 0$ is an obvious solution. We shall search for solutions $f(x) \not\equiv 0$, i.e. such that there exists a number $x_0 \in \mathbb{R}_+^*$ for which $f(x_0) \neq 0$.

Setting $y = x$ in condition (a) gives $f(xf(x)) = xf(x)$, i.e. $xf(x)$ is a fixed point for all $x \in \mathbb{R}_+^*$. To finish the problem, we must therefore determine the set F of fixed points of f.

Let y_0 be a point such that $f(y_0) = 1$. Then setting $x = x_0,\ y = y_0$ in condition (a), we find $f(x_0)(y_0 - 1) = 0 \Rightarrow y_0 = 1$. The point 1 is thus a fixed point. In fact, we shall show that this is the only fixed point.

The set F is a multiplicative subgroup of $\mathbb{R}_+^*$. That is, it has the following properties:

[1]A beautiful elementary introduction to topology may be found in [19]. In particular, Chapter 10 contains a discussion of fixed point theorems in topology.

- If $a, b \in F$, then $ab \in F$. To see this, we set $x = a, y = b$ in (a): $f(af(b)) = bf(a) \Rightarrow f(ab) = ab$, since $f(a) = a, f(b) = b$.
- If $a \in F$, then $a^{-1} \in F$. To see this, we set $x = a^{-1}, y = a$ in (a): $f(a^{-1}f(a)) = af(a^{-1}) \Rightarrow f(1) = af(a^{-1}) \Rightarrow f(a^{-1}) = a^{-1}$.

Now consider a random fixed point $a \neq 1$. Either $a > 1$ or $a < 1$. If $a > 1$ all powers a^n of a are also fixed points

$$f(a^n) = a^n .$$

In the limit $n \to \infty$, this equation gives

$$\lim_{x \to +\infty} f(x) = +\infty$$

which contradicts condition (b). Therefore, F cannot contain any fixed point strictly greater than 1. If $a < 1$, then a^{-1} is also a fixed point and so are its powers:

$$f(a^{-n}) = a^{-n} .$$

In the limit $n \to \infty$, this equation again contradicts condition (b) and F cannot contain any fixed point strictly less than 1.

Hence $xf(x) = 1$ and

$$f(x) = \frac{1}{x} . \qquad \square$$

Problem 18.2 (IMO 1994). *Let S be the set of real numbers strictly greater than -1. Find all functions $f : S \to S$ satisfying the two conditions*

(a) $f(x + f(y) + xf(y)) = y + f(x) + yf(x)$, $\forall x, y \in S$;

(b) $f(x)/x$ is strictly increasing on each of the intervals $-1 < x < 0$ and $0 < x$.

Solution. Setting $y = x$ in condition (a) gives

$$f(x + f(x) + xf(x)) = x + f(x) + xf(x) , \tag{18.1}$$

i.e. $x + f(x) + xf(x)$ is a fixed point for all $x \in S$. To finish the problem, we must therefore determine the set F of fixed points of f.

Let x_0 such that $f(x_0) = 0$. Then, in the defining equation, we set $x = y = x_0$. This gives $x_0 = 0$, i.e. 0 is a fixed point of f. In fact, we shall show that this is the only fixed point.

Now consider a random fixed point $a \neq 0$. According to equation (18.1), the points

$$\begin{aligned} a_1 &= a + f(a) + a\,f(a) = (a+1)^2 - 1 , \\ a_2 &= a_1 + f(a_1) + a_1\,f(a_1) = (a_1+1)^2 - 1 , \\ \ldots & \quad \ldots , \end{aligned}$$

are also fixed points and, in turn, the sequence

$$\frac{f(a)}{a}\,,\quad \frac{f(a_1)}{a_1}\,,\quad \frac{f(a_2)}{a_2}\,,\quad \cdots\,,$$

is the constant sequence $1,\ 1,\ 1, \cdots$. However, we notice that if $a \in (-1,0)$ then $a_1 \in (-1,0)$ and so on; if $a \in (0,+\infty)$ then $a_1 \in (0,+\infty)$ and so on. In other words, the sequence $\{f(a_n)/a_n\}$ contradicts condition (b) and therefore we cannot have a fixed point in $(-1,0) \cup (0,+\infty)$.

Hence $x + f(x) + xf(x) = 0$ which implies

$$f(x) \ = \ -\frac{x}{x+1}\,. \qquad \square$$

Problem 18.3 (IMO 1996)**.** *Let* $\mathbb{N} = \{0,1,2,3,\ldots\}$ *be the set of non-negative integers. Find all functions* f *defined on* $\mathbb{N}$ *and taking their values in* $\mathbb{N}$ *such that*

$$f(m+f(n)) \ = \ f(f(m)) + f(n)\,, \quad \textit{for all } m,n \textit{ in } \mathbb{N}\,.$$

Solution. The function $f(x) \equiv 0$ is an obvious solution. We shall search for solutions $f(x) \not\equiv 0$, i.e. such that there exists a number $n_0 \in \mathbb{N}$ for which $f(n_0) \neq 0$.

Setting $m = n = 0$ in the defining equation we find $f(f(0)) = f(f(0)) + f(0) \Rightarrow f(0) = 0$. Then, setting $m = 0$ in the defining equation, we find

$$f(f(n)) \ = \ f(n)\,. \tag{18.2}$$

Using this result, we may write the original functional equation as

$$f(m+f(n)) \ = \ f(m) + f(n)\,. \tag{18.3}$$

Equation (18.2) determines that $N = f(n)$ is a fixed points of f for all $n \in \mathbb{N}$. The set F of fixed points of F contains at least 0 and $f(n_0)$. Let's denote by N_0 the smallest non-zero integer in F. Now, let a, b be two fixed points $f(a) = a$, $f(b) = b$. Setting $m = a, n = b$ in (18.3), we find $f(a+b) = a+b$, i.e. the sum of any fixed points is also a fixed point. Therefore $k\,N_0$, for any $k \in \mathbb{N}$, is a fixed point.

If N is any fixed point, we use Euclid's algorithm to write $N = q\,N_0 + r$, with $0 \le r < N_0$. Then

$$\begin{aligned} q\,N_0 + r \ = \ N \ = \ f(N) \ &= \ f(r+qN_0) \\ &= \ f(r+f(qN_0)) && (qN_0 \text{ is a fixed point}) \\ &= \ f(r) + f(qN_0) && (\text{using equation (18.3)}) \\ &= \ f(r) + q\,N_0 && (qN_0 \text{ is a fixed point}), \end{aligned}$$

or $f(r) = r$. The residue r is a fixed point smaller than N_0. Therefore it can only be $r = 0$ and $N = q\,N_0$. The fixed points are precisely the multiples of N_0, i.e.

$$F \;=\; \{\, k\,N_0 \;,\; k \in \mathbb{N}\,\} \;.$$

Since $f(n)$ is a fixed point for any n, we conclude that $f(n)$ is a multiple of N_0 for any n. To define which multiple, we use again Euclid's algorithm to write $n = q\,N_0 + r$, $0 \leq r < N_0$. Then

$$f(n) \;=\; f(q\,N_0 + r) \;=\; f(q\,N_0) + f(r) \;=\; q\,N_0 + f(r) \;.$$

$f(r)$ must be a multiple of N_0:

$$\begin{aligned} f(0) &\;=\; 0 \;, \\ f(r) &\;=\; n_r\,N_0 \;, \quad r = 1, \ldots, N_0 - 1 \;, \end{aligned}$$

where the n_r's are $N_0 - 1$ undetermined numbers.

Therefore there is an infinite number of non-trivial solutions to the given functional equation. Each solution is characterized by a non-zero natural number N_0 and a function $\tilde{f} : \{0, 1, 2, \ldots, N_0 - 1\} \to N_0\mathbb{N}$ such that $\tilde{f}(0) = 0$. □

Part VI

GETTING ADDITIONAL EXPERIENCE

Chapter 19

Miscellaneous Problems

In this chapter we study miscellaneous problems that belong in the following categories:

- The functional equations given contain integrals (integral functional equations).
- The original problem does not refer to a functional relation but it can be solved nicely if converted to a functional relations problem.
- An assortment of some problems which do not exactly fit very well within any of the ideas previously discussed.

19.1 Integral Functional Equations

Problem 19.1. *Find the function* $f : \mathbb{R} \to \mathbb{R}$ *such that*

$$f(x) = \frac{1}{2\varepsilon} \int_{x-\varepsilon}^{x+\varepsilon} f(t)\, dt\,, \quad \forall \varepsilon > 0\,.$$

Solution. First we write:

$$\begin{aligned} f(x) &= \frac{1}{2\varepsilon} \int_{x-\varepsilon}^{x+\varepsilon} f(t)\, dt \\ &= \frac{1}{2\varepsilon} \left(\int_{0}^{x+\varepsilon} f(t)\, dt - \int_{0}^{x-\varepsilon} f(t)\, dt \right)\,. \end{aligned}$$

Differentiating with respect to t, we find

$$f'(x) = \frac{1}{2\varepsilon} \left(f(x+\varepsilon) - f(x-\varepsilon) \right)\,. \tag{19.1}$$

Since the first derivative exists, we see that also the second derivative exists:

$$f''(x) = \frac{1}{2\varepsilon} \left(f'(x+\varepsilon) - f'(x-\varepsilon) \right)\,.$$

Differentiating equation (19.1) with respect to ε:

$$0 = \frac{(f'(x+\varepsilon) + f'(x-\varepsilon))\,\varepsilon - (f(x+\varepsilon) - f(x-\varepsilon))}{2\varepsilon^2} .$$

Using the previous equation and (19.1) ,

$$f'(x+\varepsilon) + f'(x-\varepsilon) = 2\, f'(x) .$$

Differentiating with respect to ε once more:

$$f''(x+\varepsilon) = f''(x-\varepsilon) .$$

Setting $x+\varepsilon$ for x and then $x = 0$:

$$f''(2\varepsilon) = f''(0) \equiv a \Rightarrow f(y) = \frac{a}{2}\, y^2 + b\, y + c ,$$

where $y = 2\varepsilon > 0$. Substituting this result in the defining equation we find that $a = 0$. One can also show that the solution is valid when $y \leq 0$. □

Problem 19.2 (Putnam 1940)**.** *Find $f(x)$ such that*

$$\int f(x)^n\, dx = \left(\int f(x)\, dx\right)^n . \tag{19.2}$$

Comment. The statement of this problem is incomplete; no information is given about the constant n and the function $f(x)$. We shall assume that n is an integer different from 1 and that $f : \mathbb{R} \to \mathbb{R}$ is continuous.

Solution. We set

$$F(x) = \int^x f(t)\, dt \Rightarrow F'(x) = f(x) , \tag{19.3}$$

$$G(x) = \int^x f(t)^n\, dt \Rightarrow G'(x) = f(x)^n = F'(x)^n . \tag{19.4}$$

The given equation (19.2) is thus written

$$\begin{aligned} & & G(x) &= F(x)^n \\ &\overset{\text{differentiate}}{\Rightarrow} & G'(x) &= n\, F(x)^{n-1}\, F'(x) \\ &\overset{(19.4)}{\Rightarrow} & F'(x)^n &= n\, F(x)^{n-1}\, F'(x) \\ &\Rightarrow & F'(x)\left(F'(x)^{n-1} - n\, F(x)^{n-1}\right) &= 0 . \end{aligned}$$

The last equation implies that, for any x,

$$F'(x) \;=\; f(x) \;=\; 0 \quad \text{or} \quad F'(x)^{n-1} - n\,F(x)^{n-1} \;=\; 0\,.$$

The function $F'(x) = f(x) \equiv 0$ is obviously a solution. We will look for solutions that do not vanish identically. Therefore, there exists a point x_0 such that $F'(x_0) = f(x_0) \neq 0$. Since $f(x)$ is continuous, there exists a neighborhood $(x_0 - \delta, x_0 + \delta)$ (we shall determine the radius $\delta > 0$ later) of x_0 such that $F'(x) = f(x) \neq 0$ for all points x in the neighborhood. Therefore, for all points in the neighborhood

$$F'(x)^{n-1} \;=\; n\,F(x)^{n-1}\,.$$

If $n = 2k$ then

$$F'(x) \;=\; \sqrt[2k-1]{2k}\,F(x)\,,$$

and if $n = 2k + 1$ then

$$F'(x) \;=\; \pm\;\sqrt[2k]{2k+1}\,F(x)\,.$$

From the last equations we can determine the function $F(x)$ (and thus $f(x)$) easily:

$$F(x) \;=\; c\,e^{\sqrt[2k-1]{2k}\,x} \;\Rightarrow\; f(x) \;=\; c\;\sqrt[2k-1]{2k}\,e^{\sqrt[2k-1]{2k}\,x}\,,$$

and

$$F(x) \;=\; c\,e^{\pm\sqrt[2k]{2k+1}\,x} \;\Rightarrow\; f(x) \;=\; \pm\,c\;\sqrt[2k]{2k+1}\,e^{\pm\sqrt[2k]{2k+1}\,x}\,,$$

respectively. We now see that the radius of the neighborhood is infinite. □

Problem 19.3 (Putnam 1963)**.** *Find every real-valued function satisfying the condition*

$$\int_0^x f(t)dt \;=\; x\,\sqrt{f(0)f(x)}\,,\quad \forall x \geq 0\,. \tag{19.5}$$

Solution. If f is constant, $f(x) \equiv f(0)$, then

$$f(0) \;=\; |f(0)|\,,$$

which implies that $f(0) \geq 0$.

If $f(0) = 0$, then

$$\int_0^x f(t)dt \;=\; 0\,,\quad \forall x \geq 0\,.$$

This implies that $f(x) \;\equiv\; 0$.

Now let $f(0) \neq 0$. Then differentiating equation (19.5)

$$\begin{aligned} \frac{d}{dx}\int_0^x f(t)dt &= \frac{d}{dx}\left(x\sqrt{f(0)f(x)}\right) \\ \Rightarrow f(x) &= \sqrt{f(0)f(x)} + x\frac{d}{dx}\sqrt{f(0)f(x)} \\ \Rightarrow \frac{f(x)}{f(0)} &= \sqrt{\frac{f(x)}{f(0)}} + x\frac{d}{dx}\sqrt{\frac{f(x)}{f(0)}} \,. \end{aligned}$$

We define the function

$$g(x) = \sqrt{\frac{f(x)}{f(0)}} \,.$$

In terms of this function the above relation takes the form

$$g(x)^2 = g(x) + x\frac{dg(x)}{dx} \,.$$

This is a simple differential equation that can be solved by separation of variables

$$\int \frac{dx}{x} = \int \frac{dg}{g(g-1)} = \int dg\left(\frac{1}{g-1} - \frac{1}{g}\right) ,$$

or

$$\ln x = \ln|g-1| - \ln|g| - \ln c \,,$$

where c is a positive constant. The last equation is equivalent to

$$c\,x = \left|1 - \frac{1}{g}\right| \,.$$

When the expression inside the absolute value is positive, then

$$g(x) = \frac{1}{1-c\,x} \,.$$

In this case the function is defined for $x \in [0, 1/c)$. When the expression inside the absolute value is negative, then

$$g(x) = \frac{1}{1+c\,x} \,.$$

In this case the function is defined for $x \in [0, +\infty)$.

Therefore the solutions of the differential equation are

$$\begin{aligned} f(x) &= \frac{a}{(1-c\,x)^2}\,,\ a > 0, c > 0\,, \quad x \in [0, 1/c)\,, \\ f(x) &= \frac{a}{(1+c\,x)^2}\,,\ a \geq 0\,, c \geq 0\,, \quad x \in [0, +\infty)\,. \end{aligned}$$

□

Problem 19.4. *Let f be an integrable function on $(0, a)$ satisfying*

$$0 \leq f(x) \leq K \int_0^x f(t)\, dt \,,$$

for all $0 \leq x \leq a$ with $0 < K < 1$. Find all such functions.

Solution. By iteration, we find:

$$\begin{aligned}
0 \ \leq \ f(x) \ &\leq \ K \int_0^x f(t_0)\, dt_0 \\
&\leq \ K^2 \int_0^x dt_0 \int_0^{t_0} f(t_1)\, dt_1 \\
&\leq \ K^3 \int_0^x dt_0 \int_0^{t_0} dt_1 \int_0^{t_1} f(t_2)\, dt_2 \\
\cdots \\
&\leq \ K^{n+1} \int_0^x dt_0 \int_0^{t_0} dt_1 \cdots \int_0^{t_n} f(t_n)\, dt_n \,,
\end{aligned}$$

for any positive integer n. Using the well known identity

$$\int_0^x dt_0 \int_0^{t_0} dt_1 \cdots \int_0^{t_n} dt_n f(t_n) \ = \ \frac{1}{n!} \int_0^x dt (x-t)^n f(t)\, dt \,,$$

the above inequality is written as

$$0 \ \leq \ f(x) \ \leq \ \frac{1}{n!} K^{n+1} \int_0^x dt (x-t)^n f(t)\, dt \,.$$

Taking the limit of the last inequality as $n \to \infty$, we see that

$$0 \ \leq \ f(x) \ \leq 0 \,,$$

or $f(x) \equiv 0$. □

19.2 Problems Solved by Functional Relations

Perhaps the reader has met continued radicals and continued fractions. We saw one in Problem 13.4. Two additional examples are:

$$\sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2 + \ldots}}}}}}$$

and

$$\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cdots}}}}}}$$

Such expressions, for which the terms have a periodic pattern, are easy to handle. If we call them A and B respectively, we notice that they require

$$A = \sqrt{2+A}\,,$$

and

$$B = \frac{1}{2+B}\,,$$

which are easy to solve to find

$$A = 2\,, \quad B = \sqrt{2}-1\,.$$

Continued radicals and continued fractions that are not periodic are harder to handle. Also, sometimes even if an expression is periodic the algebraic equation to solve might not be very 'friendly', so we may wish to apply different methods (for such a case see Problem 19.5).

Problem 19.5 (IMO 1969 Longlist)**.** *Prove that for $a > b^2$,*

$$\sqrt{a-b\sqrt{a+b\sqrt{a-b\sqrt{a+b\sqrt{a-\ldots}}}}} = \sqrt{a-\frac{3b^2}{4}}-\frac{b}{2}\,.$$

Comment. The continued radical of the problem is obviously periodic. If we denote by C its value, then

$$C = \sqrt{a-b\sqrt{a+bC}}\,.$$

In the right hand side, C appears under two nested radicals. If we remove them, we will end up with a quartic algebraic equation. With little effort (and the known procedure for quartic equations) it can be solved but I did not try to do so. In the solution I present below, I have bypassed it.

Solution. We define

$$f(x) \equiv \sqrt{a-x\sqrt{a+x\sqrt{a-x\sqrt{a+x\sqrt{a-\ldots}}}}}\,,$$

where the domain will be set below. Obviously $f(x) > 0$ with $f(0) = \sqrt{a}$. From the definition of $f(x)$

$$f(-x) \ = \ \sqrt{a + x\sqrt{a - x\sqrt{a + x\sqrt{a - x\sqrt{a + \dots}}}}}\,,$$

and thus

$$f(-x) \ = \ \sqrt{a + x\,f(x)}\,,$$

or

$$f(-x)^2 - x\,f(x) \ = \ a\,.$$

Substituting $-x$ for x, we also have

$$f(x)^2 + x\,f(-x) \ = \ a\,.$$

Adding and subtracting the last two equations gives

$$f(-x)^2 + f(x)^2 + x\,[f(-x) - f(x)] \ = \ 2a\,, \tag{19.6}$$

$$f(x)^2 - f(-x)^2 + x\,[f(-x) + f(x)] \ = \ 0\,. \tag{19.7}$$

Equation (19.7) can be simplified as follows. First we write it in the form

$$\begin{aligned} & f(x)^2 - f(-x)^2 + x\,[f(-x) + f(x)] \ = \ 0 \\ \Rightarrow \quad & [f(x) - f(-x)]\,[f(x) + f(-x)] + x\,[f(-x) + f(x)] \ = \ 0 \\ \Rightarrow \quad & [f(x) - f(-x) + x]\,[f(x) + f(-x)] \ = \ 0\,. \end{aligned}$$

This implies that either $f(-x) + f(x) = 0$, that is f is an odd function, or $f(x) - f(-x) + x = 0$. However, the function f cannot be odd since $f(0) \neq 0$. Therefore we must necessarily have

$$f(-x) - f(x) \ = \ x\,.$$

Squaring this equation and subtracting from (19.6) also gives

$$-f(-x)\,f(x) \ = \ -a + x^2\,.$$

Since the left hand side is negative, we must restrict the domain of f to those values of x which satisfy $x^2 < a$. Looking at the last two equations, we interpret them as Vieta's formulæ for the roots $f(x)$ and $-f(-x)$ of the quadratic polynomial

$$y^2 + x\,y + (x^2 - a) \ = \ 0\,.$$

which of course we can solve easily to find

$$f(x) \ = \ -\frac{x}{2} + \sqrt{a - \frac{3x^2}{4}}\,. \qquad \square$$

Problem 19.6 (Putnam 1966)**.** *Prove*

$$\sqrt{1+2\sqrt{1+3\sqrt{1+4\sqrt{1+\dots}}}} \;=\; 3\,.$$

Comment. This problem was first posed by Ramanujan in 1911 ([6], Question No. 289). Problems for the Putman mathematical competition are not always original. Often they are known interesting ideas encountered in various areas of mathematics. To learn how Putnam problems are written read the article [54] written by a mathematician who served in the Putnam Problems Subcommittee.

Solution. We define

$$f(x) \;\equiv\; \sqrt{1+x\sqrt{1+(x+1)\sqrt{1+(x+2)\sqrt{1+\dots}}}}\,.$$

In this definition, a priori it seems that x may be any real number but this is not the case. We leave it as an exercise to our reader to discuss what happens for all values of x; we shall focus on $x \geq 1$.

From the definition

$$f(x+1) \;\equiv\; \sqrt{1+(x+1)\sqrt{1+(x+2)\sqrt{1+(x+3)\sqrt{1+\dots}}}}\,,$$

and therefore the function $f(x)$ satisfies the functional relation

$$f(x) \;=\; \sqrt{1+x\,f(x+1)}\,.$$

This may be written in the form

$$f(x+1) \;=\; \frac{f(x)^2-1}{x}\,. \tag{19.8}$$

We shall use this relation shortly.

First, we find a lower bound for $f(x)$:

$$\begin{aligned}\sqrt{1+x\sqrt{1+(x+1)\sqrt{1+(x+2)\sqrt{1+\dots}}}} \;&\geq\; \sqrt{x\sqrt{x\sqrt{x\sqrt{\dots}}}}\\ &=\; x^{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\dots}\\ &=\; x \;\geq\; \frac{x+1}{2}\,.\end{aligned}$$

Then we find an upper bound for $f(x)$:

$$\begin{aligned}
\sqrt{1+x\sqrt{1+(x+1)\sqrt{1+(x+2)\sqrt{1+\dots}}}} &\le \sqrt{(1+x)\sqrt{2(x+1)\sqrt{3(x+1)\sqrt{\dots}}}} \\
&= \sqrt{1\sqrt{2\sqrt{3\sqrt{4\dots}}}}\,(x+1)^{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\dots} \\
&< \sqrt{1\sqrt{2\sqrt{4\sqrt{8\dots}}}}\,(x+1)^{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\dots} \\
&= \sqrt{2^{\frac{1}{2}+\frac{2}{4}+\frac{3}{8}+\dots}}\,(x+1)^{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\dots} \;=\; 2(x+1)\,.
\end{aligned}$$

Therefore

$$\frac{x+1}{2} \le f(x) \le 2(x+1)\,.$$

Substituting $x+1$ for x in the above relation gives

$$\frac{x+2}{2} \le f(x+1) \le 2(x+2)\,.$$

Using equation (19.8) now, we can write it in the form

$$\frac{x(x+2)}{2} + 1 \le f(x)^2 \le 2x(x+2) + 1\,,$$

or in the stronger form

$$\frac{x(x+2)}{2} + \frac{1}{2} < f(x)^2 < 2x(x+2) + 2\,.$$

Notice that the lower bound equals $(x+1)^2/2$ and the upper bound equals $2(x+1)^2$. The last inequality implies for $f(x)$:

$$\frac{(x+1)}{\sqrt{2}} < f(x) < \sqrt{2}(x+1)\,.$$

We can repeat the above procedure a second time to find

$$\frac{(x+1)}{\sqrt[2^2]{2}} < f(x) < \sqrt[2^2]{2}(x+1)\,,$$

a third time to find

$$\frac{(x+1)}{\sqrt[2^3]{2}} < f(x) < \sqrt[2^3]{2}(x+1)\,,$$

etc, up to n times:

$$\frac{(x+1)}{\sqrt[2^n]{2}} < f(x) < \sqrt[2^n]{2}(x+1) \,,$$

In the limit $n \to +\infty$, we find

$$(x+1) \leq f(x) \leq (x+1) \;\Rightarrow\; f(x) = x+1 \,,$$

since $\sqrt[2^n]{2} \to 1$.

The number given in the problem equals $f(2) = 3$. □

Ramanujan discovered the above result while he was a student. His method then was a repeated application of the identity

$$n+2 \;=\; \sqrt{1+(n+1)(n+3)} \,.$$

That is

$$\begin{aligned} n(n+2) &= n\,\sqrt{1+(n+1)(n+3)} \\ &= n\,\sqrt{1+(n+1)\,\sqrt{1+(n+2)(n+4)}} \\ &= \ldots \end{aligned}$$

Isn't this easier than the method presented above? Of course, it is! But it ignores to verify that the passage from the finite sum to the infinite one is allowed. This is a non-trivial issue. For example, here is an alternative solution of Ramanujan's problem (taken from [48]) and which gives a different (incorrect) answer:

$$\begin{aligned} 4 &= \sqrt{1+2\,\frac{15}{2}} \\ &= \sqrt{1+2\,\sqrt{1+3\,\frac{221}{12}}} \\ &= \ldots \\ &= \sqrt{1+2\,\sqrt{1+3\,\sqrt{1+\cdots}}} \,. \end{aligned}$$

More generally, Ramanujan discovered [32] that if

$$f(x) \;\equiv\; \sqrt{ax+(n+a)^2+x\,\sqrt{a(x+n)+(n+a)^2+(x+n)\,\sqrt{a(x+2n)+(n+a)^2+(x+2n)\,\sqrt{\ldots}}}} \,,$$

then

$$f(x) = x+n+a \,.$$

Try now the following problem that was proposed by Ramanujan at the same time with the previous one:

Problem 19.7. *Evaluate the continued radical*

$$S = \sqrt{6 + 2\sqrt{7 + 3\sqrt{8 + 4\sqrt{9 + \cdots}}}}.$$

Answer. $S = 4$.

Let's return once more to the IMO 1976 problem and see one outstanding connection with continued radicals. Towards this goal, we present a beautiful formula ([37], Problem 195) that generalizes the well known results

$$\begin{aligned}
\frac{\sqrt{2}}{2} &= \sin\frac{\pi}{4}, \\
\frac{\sqrt{2-\sqrt{2}}}{2} &= \sin\frac{\pi}{8}, \\
\frac{\sqrt{2+\sqrt{2}}}{2} &= \cos\frac{\pi}{8} = \sin\frac{3\pi}{8}, \\
&\ldots
\end{aligned}$$

Lemma 19.1. *For $\alpha_i \in \{-1, 1\}$, $i = 0, 1, \ldots, n$,*

$$2\sin\left[\left(\alpha_0 + \frac{\alpha_0\alpha_1}{2} + \cdots + \frac{\alpha_0\alpha_1\ldots\alpha_n}{2^n}\right)\frac{\pi}{4}\right] = \alpha_0\sqrt{2 + \alpha_1\sqrt{2 + \alpha_2\sqrt{\cdots + \alpha_n\sqrt{2}}}}.$$

The proof is easy by induction. Given the continued radical,

$$\alpha_0\sqrt{2 + \alpha_1\sqrt{2 + \alpha_2\sqrt{2 + \ldots}}},$$

according to this lemma, its partial sums are given by

$$x_n = 2\sin\left[\left(\alpha_0 + \frac{\alpha_0\alpha_1}{2} + \cdots + \frac{\alpha_0\alpha_1\ldots\alpha_{n-1}}{2^{n-1}}\right)\frac{\pi}{4}\right].$$

The series

$$\alpha_0 + \frac{\alpha_0\alpha_1}{2} + \cdots + \frac{\alpha_0\alpha_1\ldots\alpha_{n-1}}{2^{n-1}} + \cdots$$

is absolutely convergent (since the absolute value of the summands are the terms of a geometric progression with ratio 1/2) and thus it converges to some number α. Therefore the original continued radical converges to the real number

$$x = 2\sin\frac{\alpha\pi}{4}.$$

Problem 19.8 (IMO 1976). *Let $f(x) = x^2 - 2$ and $f^j(x) = f(f^{j-1}(x))$ for $j = 2, 3, \cdots$. Show that for any positive integer n, the roots of the equation $f^n(x) = x$ are real and distinct.*

Solution. We observe that the equation $f^j(x) = x$ can be written $f(f^{j-1}(x)) = x$ or $f^{j-1}(x)^2 - 2 = x$ and from this

$$f^{j-1}(x) = \pm \sqrt{2+x}\,.$$

By repeating this process,

$$f^{j-2}(x) = \pm \sqrt{2 \pm \sqrt{2+x}}\,,$$

and inductively after n steps, starting with $j = n$,

$$x = \pm \sqrt{2 \pm \sqrt{2 \pm \sqrt{2 \pm \ldots \sqrt{2+x}}}}\,,$$

where we have n nested radicals. There are 2^n choices in the signs. Given a combination of signs $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$,

$$x = \alpha_0 \sqrt{2 + \alpha_1 \sqrt{2 + \alpha_2 \sqrt{2 + \cdots + \alpha_{n-1} \sqrt{2+x}}}}\,,$$

by repetitive substitution of the x in the right hand side, x is given by a periodic continued radical

$$x = \alpha_0 \sqrt{2 + \alpha_1 \sqrt{2 + \alpha_2 \sqrt{2 + \ldots}}}\,, \quad \alpha_{n+k} = \alpha_k\,, \quad k = 0, 1, \ldots, n-1\,.$$

This corresponds to a real number $2\sin(\alpha\pi/4)$. Therefore, each choice in the combination of signs corresponds to one root; there are 2^n real solutions. □

To understand how this works and compare with the result of Problem 1.6, in Table 19.1 I have listed the roots of $f^3(x) = x$ and the corresponding choices of signs. In doing so, I have omitted the calculations assuming that you can carry out the (easy) summation of the series.

The following problem I wrote down in direct analogy with Ramanujan's continued radical.

Problem 19.9. *Evaluate the continued fraction*

$$\cfrac{1}{2 + \cfrac{1}{3 + \cfrac{1}{4 + \cfrac{1}{5 + \cfrac{1}{6 + \cfrac{1}{7 + \cdots}}}}}}\,.$$

α_0	α_1	α_2	$\sin\frac{\alpha\pi}{4}$	$\cos\left(\frac{\pi}{2}-\frac{\alpha\pi}{2}\right)$
-1	-1	-1	$\sin\left(-\frac{\pi}{6}\right)$	$\cos\frac{2\pi}{3}$
-1	-1	$+1$	$\sin\left(-\frac{\pi}{14}\right)$	$\cos\frac{4\pi}{7}$
-1	$+1$	-1	$\sin\left(-\frac{5\pi}{14}\right)$	$\cos\frac{6\pi}{7}$
-1	$+1$	$+1$	$\sin\left(-\frac{7\pi}{18}\right)$	$\cos\frac{8\pi}{9}$
$+1$	-1	-1	$\sin\left(\frac{3\pi}{14}\right)$	$\cos\frac{2\pi}{7}$
$+1$	-1	$+1$	$\sin\left(\frac{\pi}{18}\right)$	$\cos\frac{4\pi}{9}$
$+1$	$+1$	-1	$\sin\left(\frac{5\pi}{18}\right)$	$\cos\frac{2\pi}{9}$
$+1$	$+1$	$+1$	$\sin\left(\frac{\pi}{2}\right)$	$\cos 0$

Table 19.1: Roots of $f^3(x) = x$ using the method of continued radicals. The last column gives the roots as the cosine of an angle. These answer is exactly that of Problem 1.6. Also, we see that when the signs are chosen such that $\alpha_0\alpha_1\alpha_2 = 1$, the continued radical provides the roots corresponding to the angles $\theta_k = \frac{2\pi k}{2^3-1}$ and when the signs are chosen such that $\alpha_0\alpha_1\alpha_2 = -1$, the continued radical provides the roots corresponding to the angles $\theta_k = \frac{2\pi k}{2^3+1}$.

Solution. In order to better understand the pattern, we will define the function

$$f(x) = \cfrac{1}{x + \cfrac{1}{(x+1) + \cfrac{1}{(x+2) + \cfrac{1}{(x+3) + \dots}}}} = \cfrac{1}{1 + \cfrac{\frac{1}{x(x+1)}}{1 + \cfrac{\frac{1}{(x+1)(x+2)}}{1 + \cfrac{\frac{1}{(x+2)(x+3)}}{1 + \dots}}}}.$$

Imagine that we have found functions $f_n(x)$, $n = 0, 1, 2, \dots$ and $k_n(x)$ $n = 1, 2, \dots$ such that

$$f_{n-1} - f_n = k_n f_{n+1}, \quad n = 1, 2, \dots .$$

Then

$$\begin{aligned}
\frac{f_{n-1}}{f_n} - 1 &= k_n \frac{f_{n+1}}{f_n} \\
\Rightarrow \quad \frac{f_{n-1}}{f_n} &= 1 + k_n \frac{f_{n+1}}{f_n} \\
\Rightarrow \quad \frac{f_n}{f_{n-1}} &= \frac{1}{1 + k_n \frac{f_{n+1}}{f_n}}.
\end{aligned}$$

Applying this equation repeatedly

$$\frac{f_1}{f_0} = \frac{1}{1+k_1\frac{f_2}{f_1}} = \frac{1}{1+k_1\cfrac{1}{1+k_2\frac{f_3}{f_2}}} = \frac{1}{1+k_1\cfrac{1}{1+k_2\cfrac{1}{1+k_3\frac{f_4}{f_3}}}} = \frac{1}{1+\cfrac{k_1}{1+\cfrac{k_2}{1+\cfrac{k_3}{1+\cdots}}}}\,.$$

Therefore it remains to find the functions f_n such that

$$f_{n-1} - f_n = \frac{1}{(x+n-1)(x+n)}\, f_{n+1}\,.$$

With a little of effort (trial and error), one can discover that

$$f_0(x) = 1 + \frac{1}{x}\frac{1}{1!} + \frac{1}{x(x+1)}\frac{1}{2!} + \frac{1}{x(x+1)(x+3)}\frac{1}{3!} + \cdots\,,$$

and $f_n(x) = f_{n-1}(x+1)$, $n = 1, 2, 3, \ldots$ Then

$$f(x) = \frac{f_0(x+1)}{f_0(x)}\,.$$

In particular, the continued fraction of the problem has the value

$$f(2) = \frac{f_0(3)}{f_0(2)}\,,$$

with

$$f_0(2) = \sum_{k=0}^{\infty}\frac{1}{k!}\frac{1}{(k+1)!}\,, \quad f_0(3) = 2\sum_{k=0}^{\infty}\frac{1}{k!}\frac{1}{(k+2)!}\,. \qquad \square$$

You might be bothered by the fact that the final result is still in terms of infinite series. Well, you are not alone in this. However, this problem, besides teaching you a way to use functions in deriving results for continued fractions, it should also be used as a lesson that often in mathematics results in closed form may not be possible. Instead, we present them in terms of well known special functions. Readers with experience in such a topic might have recognized the special function

$${}_0F_1(b,z) = 1 + \frac{1}{b}\frac{z}{1!} + \frac{1}{b(b+1)}\frac{z^2}{2!} + \frac{1}{b(b+1)(b+2)}\frac{z^3}{3!} + \cdots\,,$$

which is known as the (0,1) **hypergeometric function**. The indices 0 and 1 count the number of parameters in the numerator and denominator of each term (no parameter and one parameter respectively for ${}_0F_1$). In general, one can define the ${}_nF_m(a_1,\ldots,a_n,b_1,\ldots,b_m,z)$

with n parameters in the numerator and m parameters in the denominator of each term. For example, the ${}_1F_1$ is:

$$ {}_1F_1(a,b,z) \;=\; 1+\frac{a}{b}\,\frac{z}{1!}+\frac{a(a+1)}{b(b+1)}\,\frac{z^2}{2!}+\frac{a(a+1)(a+2)}{b(b+1)(b+2)}\,\frac{z^3}{3!}+\cdots\,. $$

There is an extensive theory for hypergeometric functions. In particular, the method I used in the solution of the problem is due to Gauss who studied the hypergeometric functions systematically. If you wish to gain additional experience, try the following problem.

Problem 19.10. *Prove that*

$$ \tanh z \;=\; \cfrac{z}{1+\cfrac{z^2}{3+\cfrac{z^2}{5+\cfrac{z^2}{7+\cdots}}}}\,. $$

Hint. It is true that

$$ \cosh z \;=\; {}_0F_1\left(\frac{1}{2},\frac{z^2}{4}\right)\,, \quad \frac{\sinh z}{z} \;=\; {}_0F_1\left(\frac{3}{2},\frac{z^2}{4}\right)\,. $$ □

In Chapter 15 we mentioned that convex and concave functions are useful for proving inequalities although we did not elaborate extensively. Functions and appropriate functional relations (beyond the Jensen inequality) can be very helpful in proving inequalities. Such is the following example.

Problem 19.11 (BWMC 1992)**.** *Let* $a=\sqrt[1992]{1992}$. *Which of the following two numbers is greater?*

$$ 1992\,, \quad a^{a^{a^{\cdot^{\cdot^{\cdot^{a}}}}}}\,, $$

where the a appears 1992 times in the second expression.

Solution. We will consider the function $f(x)=a^x$ since

$$ f^n(x) \;=\; a^{a^{a^{\cdot^{\cdot^{\cdot^{a^x}}}}}}\,, $$

where the a appears n times. Notice that $x=1992$ is a fixed point of $f(x)$

$$ f(1992) \;=\; 1992\,, $$

and therefore $f^n(1992)=1992$ for all n as we can easily verify.

Then we notice that

$$a \;=\; \sqrt[1992]{1992} \;=\; e^{\frac{\ln 1992}{1992}} > 1 + \frac{\ln 1992}{1992} > 1 \,,$$

where we used the exponential inequality $e^x > 1 + x$ for $x \neq 0$. And since $a > 1$, the function $f(x) = a^x$ is strictly increasing. For $x > y$ it is then $f(x) > f(y)$. By repeated application of this relation,

$$x > y \;\Rightarrow\; f^n(x) > f^n(y) \,.$$

We now apply the last functional relation for $x = 1992$, $y = a$, and $n = 1991$:

$$f^{1991}(1992) > f^{1991}(a) \;\Rightarrow\; 1992 > a^{a^{a^{\cdot^{\cdot^{\cdot^{a}}}}}} \,. \qquad \square$$

19.3 Assortment of Problems

Problem 19.12 (IMO 1990)**.** *Construct a function $f : \mathbb{Q}_+^* \to \mathbb{Q}_+^*$ such that*

$$f(xf(y)) \;=\; \frac{f(x)}{y} \,, \quad \forall x, y \in \mathbb{Q}_+^* \,. \tag{19.9}$$

Solution. Let $x = 1/f(y)$. Then

$$f\left(\frac{1}{f(y)}\right) \;=\; y\, f(1) \,. \tag{19.10}$$

From this equation we see that f is injective:

$$f(a) \;=\; f(b) \;\Rightarrow\; f\left(\frac{1}{f(a)}\right) \;=\; f\left(\frac{1}{f(b)}\right) \;\Rightarrow\; a\, f(1) \;=\; b\, f(1) \;\Rightarrow\; a \;=\; b \,.$$

Next let $y = 1$ in (19.9):

$$f(xf(1)) \;=\; f(x) \,.$$

Since f is , we conclude that $x f(1) = x$, or[1]

$$f(1) \;=\; 1 \,.$$

Setting $x = 1$ now in the defining equation (19.9), we find

$$f(f(y)) \;=\; \frac{1}{y} \,. \tag{19.11}$$

Finally, we can prove that f is multiplicative. Replacing y by $1/f(y)$ in (19.9) and making use of (19.10), we find that f satisfies the Cauchy IV equation:

$$f(x\,y) \;=\; f(x)\, f(y) \,.$$

[1]Apparently, the point $x = 1$ is a fixed point but this interpretation is not important in this problem.

Any $x \in \mathbb{Q}_+^*$ can be uniquely expanded in a product of primes

$$x = \prod_i p_i^{n_i},$$

where $n_i \in \mathbb{Z}^*$. From the Cauchy IV equation

$$f(x) = f(\prod_i p_i^{n_i}) = \prod_i (f(p_i))^{n_i}.$$

Applying once more

$$f(f(x)) = f\left(\prod_i (f(p_i))^{n_i}\right) \overset{(19.11)}{\Rightarrow} \frac{1}{x} = \prod_i (f(f(p_i)))^{n_i},$$

or

$$\prod_i (f(f(p_i)))^{n_i} = \prod_i p_i^{-n_i}.$$

The two sides will agree for any function f whose values at the prime numbers are chosen arbitrarily, but otherwise to satisfy

$$f(f(p_i)) = \frac{1}{p_i}, \quad \forall p_i.$$

For example, we can partition the set of primes into two sets Q and Q' and order their elements: $Q = \{q_1, q_2, q_3, \ldots, ...\}$ and $Q' = \{q'_1, q'_2, q'_3, \ldots, ...\}$. Then define f such that $f(q_i) = q'_i$ and $f(q'_i) = 1/q_i$, $\forall i$. □

Problem 19.13 (IMO 1986)**.** *: Find all functions $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that*
(a) $f(2) = 0$,
(b) $f(x) \neq 0$ *for* $0 \leq x < 2$,
(c) $f(xf(y))\, f(y) = f(x+y)$, $\forall x, y \in \mathbb{R}_+$.

Solution. Setting $y = 2$ in (c) and using (a) we find

$$f(x+2) = f(xf(2))\, f(2) = 0.$$

Since $x \geq 0$ in (c), this result implies that

$$f(x) = 0, \quad \forall x \geq 2.$$

Setting $x = y = 0$ in (c) we find that $f(0) = 0, 1$. But $f(0) \neq 0$ from (b) and hence $f(0) = 1$.

Now, set $x = 2 - y$, $0 < y < 2$ in (c):

$$f((2-y)f(y))\,f(y) \;=\; f(2) \;=\; 0\,.$$

Since the condition (b) is given, $f((2-y)f(y)) = 0$. This necessarily implies that $(2-y)\,f(y) \geq 2$ or

$$f(y) \geq \frac{2}{2-y}\,, \quad 0 < y < 2\,. \tag{19.12}$$

In fact, we will show now that

$$f(x) \;=\; \frac{2}{2-x}\,, \quad 0 < x < 2\,.$$

Assume that the above is not the case. Then there are points in $(0,2)$ which satisfy the strict inequality of (19.12). Let x_0 be the *first* point in $(0,2)$ such that $f(x_0) > 2/(2-x_0)$. Then define $x_0' \neq x_0$ such that

$$f(x_0') \;=\; \frac{2}{2-x_0}\,.$$

If $x_0' < x_0$ we must have

$$f(x_0') \;=\; \frac{2}{2-x_0'}\,.$$

This, in conjunction with the definition of x_0' implies that $x_0 = x_0'$ which is not possible. Therefore, we must assume that $x_0' > x_0$. Now let $x_0'' = 2 - x_0'$. The condition $x_0' > x_0$ gives $2 - x_0' < 2 - x_0 \Rightarrow x_0'' < 2 - x_0$ or

$$x_0'' + x_0 \;<\; 2\,.$$

In (c), set $x = x_0''$ and $y = x_0$:

$$f(x_0''\,f(x_0')) \;=\; f(x_0'' + x_0') \;=\; f(2) \;=\; 0 \;\Rightarrow\; x_0''\,f(x_0') \geq 2 \;\Rightarrow\; x_0''\,\frac{2}{2-x_0} \geq 2\,,$$

or

$$x_0'' + x_0 \geq 2\,.$$

However, this inequality contradicts the previous one. So we must have

$$f(x) \;=\; \begin{cases} \frac{2}{2-x}\,, & \text{if } x \in [0,2)\,, \\ 0\,, & \text{if } x \in [2,+\infty)\,. \end{cases}$$ □

Problem 19.14 (Greece 1996). *Let $f : (0,+\infty) \to \mathbb{R}$ be a function such that*
(a) f is strictly increasing;
(b) $f(x) > -\frac{1}{x}$, $\forall x > 0$;
(c) $f(x)\,f(f(x) + \frac{1}{x}) = 1$, $\forall x > 0$.
Find $f(1)$.

Solution. From (c) we see immediately that $f(x) \neq 0$, $\forall x > 0$. Now let $y = f(x) + \frac{1}{x} > 0$ for $x > 0$. In (b) we substitute x with y:

$$f(y) > -\frac{1}{y}\,.$$

At the same (c) gives

$$f(x) \;=\; \frac{1}{f(y)}\,. \tag{19.13}$$

Also, in (c), we substitute x with y:

$$f(y)\,f\left(f(y)+\frac{1}{y}\right) \;=\; 1 \;\Rightarrow\; \frac{1}{f(y)} \;=\; f\left(f(y)+\frac{1}{y}\right) \overset{(19.13)}{\Rightarrow} f(x) \;=\; f\left(\frac{1}{f(x)}+\frac{1}{y}\right)\,.$$

Since f is strictly increasing, it is injective, and therefore $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$. From the equation we found above we thus conclude that

$$x \;=\; \frac{1}{f(x)} + \frac{1}{y}\,,$$

or

$$x \;=\; \frac{1}{f(x)} + \frac{1}{f(x)+\frac{1}{x}}\,.$$

We can easily rewrite this equation as a quadratic one in $f(x)$:

$$x^2 f(x)^2 - x f(x) - 1 \;=\; 0\,,$$

with solutions

$$f(x) \;=\; \frac{\phi}{x}\,, \quad f(x) \;=\; -\frac{1}{\phi\, x}\,,$$

where

$$\phi \;=\; \frac{1+\sqrt{5}}{2}$$

is the golden ratio. Only one solution is strictly increasing:

$$f(x) \;=\; -\frac{1}{\phi\, x}\,.$$

So finally $f(1) = -1/\phi$. □

Chapter 20

Unsolved Problems

A note of warning for the reader: The separation of the problems to categories in this chapter is artificial. I only do it, in case one wishes to find and work on a group of problems containing (but not focusing exclusively) on a particular idea.

20.1 Functions

Problem 20.1. *Show that the function $f : \mathbb{R} \to \mathbb{R}$ defined by*

$$f(x) = \lim_{n\to\infty} \lim_{m\to\infty} \cos(n!\pi x)^{2m}\,, \quad n, m \in \mathbb{N}^*\,,$$

is the Dirichlet function.

Problem 20.2 (Thomæ's ruler function)**.** *If $x \in \mathbb{Q}$, let $x = m/n$ where $n > 0$ and m, n have no common divisor greater than 1. Then we define*

$$f(x) = \begin{cases} \frac{1}{n}\,, & x \in \mathbb{Q}\,, \\ 0\,, & x \in \mathbb{R} \setminus \mathbb{Q}\,. \end{cases}$$

Show that f is discontinuous at every rational number in [0,1] and continuous at every irrational number in [0,1].

Problem 20.3 ([58])**.** *Let $g(x)$ be a function with domain [0,2] defined as follows:*

$$g(x) = \begin{cases} 0\,, & x \in [0, \frac{1}{3}] \cup [\frac{5}{3}, 2], \\ 3x - 1\,, & x \in [\frac{1}{3}, \frac{2}{3}], \\ 1\,, & x \in [\frac{2}{3}, \frac{4}{3}], \\ -3x + 5\,, & x \in [\frac{4}{3}, \frac{5}{3}]. \end{cases}$$

We then define the extension of g to all $\mathbb{R}$ by

$$g(x+2) = g(x)\,.$$

Finally we define the function $f:[0,1]\to\mathbb{R}^2$ by $t\mapsto(x,y)$ such that

$$x(t) = \sum_{n=1}^{\infty}\frac{1}{2^n}\,g\Big(3^{2n-2}t\Big)\,,$$
$$y(t) = \sum_{n=1}^{\infty}\frac{1}{2^n}\,g\Big(3^{2n-1}t\Big)\,.$$

Prove that the $f([0,1]) = [0,1]\times[0,1]$.

Comment. A curve such as the one given in the above problem is called a **space filling curve**. Peano was the first to consider such curves.

Problem 20.4. *Prove that there is no continuous injective function $f:[0,1]\to[0,1]\times[0,1]$ such that $f([0,1]) = [0,1]\times[0,1]$.*

Problem 20.5 ([11], Problem 4.17)**.** *The function $f\big|[0,1]$ is defined as follows:*

$$f(x) = \begin{cases} x\,, & x = \text{rational}\,,\\ 1-x\,, & x = \text{irrational}\,.\end{cases}$$

Prove that:

(a) $f^2(x) = x$, for all $x\in[0,1]$.
(b) $f(x)+f(1-x) = 1$, for all $x\in[0,1]$.
(c) f is continuous only at $x = 1/2$.
(d) f assumes every value between 0 and 1.
(e) $f(x+y)-f(x)-f(y)$ is rational for all $x,y\in[0,1]$.

Problem 20.6. *Let a function $f:\mathbb{R}\to\mathbb{R}$ that satisfies the Cauchy II equation*

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

and for which there is a function $g:\mathbb{R}\to\mathbb{R}$ such that

$$f(x) = 1 + x\,g(x)\,,\quad \forall x,y\in\mathbb{R}\,.$$

If

$$\lim_{x\to 0} g(x) \;=\; 1\,,$$

show that the derivative $f'(x)$ *exists for all* $x \in \mathbb{R}$, *and that*

$$f'(x) \;=\; f(x)\,.$$

Problem 20.7 ([4], Problem 1795)**.** *Find a function* $f : [0,1] \to [0,1]$ *such that for each non-trivial interval* $I \subseteq [0,1]$, *we have* $f(I) = [0,1]$.

Problem 20.8 ([2], Problem 886)**.** *Call a function* f ***good*** *if* $f^{2008}(x) = -x$ *for all* $x \in \mathbb{R}$, *where* f^{2008} *denotes the 2008-th iterate of* f. *Prove the following:*

(a) Every good function is bijective and odd and cannot be monotonic.

(b) If f *is good and* $x_0 \neq 0$, *there exist infinitely many 5-tuples* $(p_1, p_2, p_3, p_4, p_5)$ *of distinct positive integers whose sum is a multiple of 5 and for which, with* $q_k = f^{p_k}(x_0)$, $q_1 \neq q_i$ *for* $i = 2, 3, 4, 5$ *and* $q_i \neq q_{i+1}$ *for* $i = 2, 3, 4$.

Problem 20.9 (Greece 1983)**.** *If the function* $f : [0, +\infty) \to \mathbb{R}$ *satisfies the relation*

$$f(x)\,e^{f(x)} \;=\; x\,,$$

for all x *in its domain, prove that:*
(a) f *is monotonic over its entire domain;*
(b) $\lim\limits_{x\to+\infty} f(x) = +\infty$;
(c) $\lim\limits_{x\to+\infty} \frac{f(x)}{\ln x} = 1$.

Problem 20.10 (Greece 1983)**.** *The function* $f : \mathbb{R} \to \mathbb{R}$ *is defined by* $f(x) = x^5 + x - 1$.
(a) Prove that f *is bijective.*
(b) Show that $f(1001^{999}) < f(1001^{1000})$.
(c) Determine the roots of the equation $f(x) = f^{-1}(x)$.

Problem 20.11 (LMO 1989). *How many real solutions does the following equation have?*

$$\sin(\sin(\sin(\sin(\sin(x))))) = \frac{x}{3}$$

Problem 20.12 (Putnam 1998). *Let f be a real function on the real line with continuous third derivative. Prove that there exists a point a such that*

$$f(a) \cdot f'(a) \cdot f''(a) \cdot f'''(a) \geq 0 .$$

Problem 20.13 (Russia 1998). *Let $f(x) = x^2 + ax + b\cos x$. Find all values of a, b for which the equations $f(x) = 0$ and $f(f(x)) = 0$ have the same (non-empty) set of real roots.*

Problem 20.14 (Mongolia 2000). *A function $f : \mathbb{R} \to \mathbb{R}$ satisfies the following conditions:*
(a) $|f(a) - f(b)| \leq |a - b|$ for any real numbers $a, b \in \mathbb{R}$.
(b) $f^3(0) = 0$.
Prove that $f(0) = 0$.

Problem 20.15 (IMO 1970 Longlist). *Suppose that f is a real function defined for $0 \leq x \leq 1$ having the first derivative f' for $0 \leq x \leq 1$ and the second derivative f'' for $0 < x < 1$. Prove that if*

$$f(0) = f'(0) = f'(1) = f(1) - 1 = 0 ,$$

there exists a number $0 < y < 1$ such that $|f''(y)| \geq 4$.

Problem 20.16 (Iran 1997). *Suppose that $f : \mathbb{R} \to \mathbb{R}$ has the following properties:*
(a) $f(x) \leq 1$, $\forall x \in \mathbb{R}$,
(b) $f\left(x + \frac{13}{42}\right) + f(x) = f\left(x + \frac{1}{6}\right) + f\left(x + \frac{1}{7}\right)$, $\forall x \in \mathbb{R}$.
Prove that f is periodic.

Problem 20.17 (IMO 1978 Longlist). *Assume that the functions* $f_1, f_2, \ldots, f_n : I \to \mathbb{R}_+$ *are concave. Prove that their geometric mean*

$$f = \sqrt[n]{f_1 f_2 \cdots f_n}$$

is also a concave function.

Problem 20.18 (Turkey 1998). *Let* $A = \{1, 2, 3, 4, 5\}$. *Find the number of functions from the set of non-empty subsets of* A *to* A *for which* $f(B) \in B$ *for every* $B \subseteq A$ *and* $f(B \cup C) \in \{f(B), f(C)\}$ *for every* $B, C \subseteq A$.

Problem 20.19 ([1], Problem 11472). *Let* n *be a non-negative integer, and let* f *be a* $(4n+3)$*-times continuously differentiable function on* $\mathbb{R}$. *Show that there is a number* ξ *such that at* $x = \xi$,

$$\prod_{k=0}^{4n+3} \frac{d^k f(x)}{dx^k} \geq 0\,.$$

Problem 20.20 ([1], Problem 11215). *A car moves along the real line from* $x = 0$ *at* $t = 0$ *to* $x = 1$ *at* $t = 1$, *with differentiable position function* $x(t)$ *and differentiable velocity function* $v(t) = x'(t)$. *The car begins and ends the trip at a standstill; that is,* $v = 0$ *at both the beginning and the end of the trip. Let* V *be the maximum velocity attained during the trip. Prove that at some time between the beginning and end of the trip,* $|v'| > V^2/(V-1)$.

Problem 20.21 ([1], Problem 11221). *Give an example of a function* g *from* $\mathbb{R}$ *into* $\mathbb{R}$ *such that* g *is differentiable everywhere,* g' *is differentiable on one dense subset of* $\mathbb{R}$, *and* g' *is discontinuous on another dense subset of* $\mathbb{R}$.

Problem 20.22 ([1], Problem 11232). *Let* f *be a continuous function from* $\mathbb{R}$ *into* $\mathbb{R}$ *that is lower bounded. Show that there exists a real number* x_0 *such that* $f(x_0) - f(x) < |x - x_0|$ *holds for all* x *other than* x_0.

Problem 20.23 ([4], Problem 1732). *For each positive integer n define the function $f_n : [0,1]^n \to [0,1]^2$ by*

$$f(x_1, x_2, \ldots, x_n) \;=\; \left(\frac{x_1 + x_2 + \cdots + x_n}{n}, \sqrt[n]{x_1 x_2 \ldots x_n}\right),$$

and let $I(n)$ be the image of f_n. Determine

$$\bigcup_{n=1}^{\infty} I(n) \,.$$

Problem 20.24 (Putnam 1974). *In the standard definition, a real-valued function of two real variables $g : \mathbb{R}^2 \to \mathbb{R}$ is continuous if, for every point $(x_0, y_0) \in \mathbb{R}^2$ and every $\varepsilon > 0$, there is a $\delta > 0$ such that $\sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$ implies $|g(x, y) - g(x_0, y_0)| < \varepsilon$.*

By contrast, $f : \mathbb{R}^2 \to \mathbb{R}$ is said to be continuous in each variable separately if, for each fixed value y_0 of y, the function $f(x, y_0)$ is continuous in the usual sense as a function of x, and similarly $f(x_0, y)$ is continuous as a function of y for each fixed x_0.

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be continuous in each variable separately. Show that there exists a sequence of continuous functions $g_n : \mathbb{R}^2 \to \mathbb{R}$ such that

$$f(x, y) \;=\; \lim_{n\to\infty} g_n(x, y)\,, \quad \forall (x, y) \in \mathbb{R}^2 \,.$$

Problem 20.25 ([1], Problem 11009). *Consider the function f defined by*

$$f(x) \;=\; \frac{a^x - b^x}{c^x - d^x}\,,$$

where $a > b \geq c > d > 0$, with the removable singularity at zero filled in. Prove that f is convex on $\mathbb{R}$, and that $\log f$ is either convex or concave on $\mathbb{R}$. Determine the values of a, b, c, d for which $\log f$ is convex on $\mathbb{R}$.

Problem 20.26 ([3], Problem 2700). *Show that the function e^{-xn^2} can be written in the following form:*

$$e^{-xn^2} \;=\; \sum_{k=0}^{n-1} (-1)^k \frac{n^{2k} x^k}{k!} + (-1)^n \frac{n^{2n} x^n}{n!}\,\phi_x(n)\,,$$

where

$$\phi_x(n) \; = \; 1 - \int_0^{xn^2} e^{-t} \left(1 - \frac{t}{xn^2}\right)^n \, dt \, .$$

Determine the leading large n behavior of $\phi_x(n)$, and show that

$$\lim_{n\to\infty} n\phi_x(n) \; = \; \frac{1}{x} \, .$$

Problem 20.27 (Ukraine 1998)**.** *Let a function $f : [0,1] \to [0,1]$ such that*

$$f\Big(f(x) + y\Big) \; = \; f(x) + f(y) \, ,$$

for all $x, y \in [0,1]$. It is also known that there is a $\lambda \in (0,1)$ such that $f(\lambda) \neq 0, \lambda$.

(a) Give an example of such a function.

(b) Prove that for any $x \in [0,1]$,

$$f^{19}(x) \; = \; f^{98}(x) \, .$$

Problem 20.28 (HIMC 1999)**.** *The function*

$$f(x, y, z) \; = \; \frac{x^2 + y^2 + z^2}{x + y + z}$$

is defined for every x, y, z such that $x + y + z \neq 0$. Find a point (x_0, y_0, z_0) such that

$$0 < x_0^2 + y_0^2 + z_0^2 < \frac{1}{1999}$$

and

$$1.999 < f(x_0, y_0, z_0) < 2 \, .$$

20.2 Problems That Can Be Solved Using Functions

Problem 20.29 (Bulgaria 2000)**.** *Prove that for any two real numbers a and b there exists a real number $c \in (0,1)$ such that*

$$\left| ac + b + \frac{1}{c+1} \right| > \frac{1}{24} \, .$$

Problem 20.30 (ROM 1996)**.** *Prove the inequality*

$$2^{a_1} + 2^{a_2} + \cdots + 2^{a_{1996}} \leq 1995 + 2^{a_1 + a_2 + \cdots + a_{1996}} \, ,$$

for any real non-positive numbers $a_1, a_2, \ldots, a_{1996}$.

20.3 Arithmetic Functions

Problem 20.31. *Let a sequence of real numbers a_n satisfying the condition*

$$a_n = \lambda\, a_{n-1} + \omega\, n\mu^n\ , \quad n = 1, 2, \dots\ ,$$

with $\lambda \neq 0$. Find the expression of a_n in terms of a_0.

Problem 20.32 (Putnam 1963)**.** *Let f be a function with the following properties:*
(a) $f(n)$ is defined for every positive integer n;
(b) $f(n)$ is a positive integer;
(c) $f(2) = 2$;
(d) $f(mn) = f(m)\, f(n)$, for all relatively prime integers m and n;
(e) f is strictly increasing.
Prove that $f(n) = n$ for $n = 1, 2, 3 \dots$.

Problem 20.33 (Australia 1984)**.** *A non-negative integer $f(n)$ is assigned to each positive integer n in such a way that the following conditions are satisfied:*
(a) $f(mn) = f(m) + f(n)$, for all positive integers m, n;
(b) $f(n) = 0$, whenever the final (right-hand) decimal digit of n is 3;
(c) $f(10) = 0$.
Calculate $f(1984)$ and $f(1985)$.

Problem 20.34 (Austria 1986)**.** *For n a given positive integer, determine all functions $F : \mathbb{N} \to \mathbb{R}$ such that*

$$F(x + y) \;=\; F(xy - n)\ , \quad \forall x, y \in \mathbb{N}\ , \quad xy > n\ .$$

Problem 20.35 (Estonia 2000; Netherlands TST 2008)**.** *Find all functions $f : \mathbb{N}^* \to \mathbb{N}^*$ such that*

$$f^3(n) + f^2(n) + f(n) \;=\; 3n\ ,$$

for all $n \in \mathbb{N}^*$.

Problem 20.36 (BMO 2009)**.** *Denote by $\mathbb{N}^*$ the set of all positive integers. Find all functions $f : \mathbb{N}^* \to \mathbb{N}^*$ such that*

$$f\left(f^2(m) + 2f^2(n)\right) = m^2 + 2n^2 ,$$

for all $m, n \in \mathbb{N}^*$.

Problem 20.37 (IMO 1988)**.** *A function f is defined on the positive integers by*

$$\begin{aligned} f(1) &= 1 , \quad f(3) = 3 , \\ f(2n) &= f(n) , \\ f(4n+1) &= 2f(2n+1) - f(n) , \\ f(4n+3) &= 3f(2n+1) - 2f(n) , \end{aligned}$$

for all positive integers n. Determine the number of positive integers n, less than or equal to 1988, for which $f(n) = n$.

Problem 20.38 (IMO 1988 Shortlist)**.** *Let $f(n)$ be a function defined on the set of all positive integers and having its values in the same set. Suppose that*

$$f(f(n) + f(m)) = m + n ,$$

for all positive integers n, m. Find the possible value for $f(1988)$.

Problem 20.39 (IMO 1998)**.** *Consider all functions f from the set $\mathbb{N}^*$ of all positive integers into itself satisfying the equation*

$$f\left(t^2 f(s)\right) = f(t)^2 s ,$$

for all s and t in $\mathbb{N}^*$. Determine the least possible value for $f(1998)$.

Problem 20.40 (Putnam 1957)**.** *If facilities for division are not available, it is sometimes convenient in determining the decimal expansion of $1/A$, $A > 0$ to use the iteration*

$$X_{k+1} = X_k (2 - AX_k) , \quad k = 0, 1, 2, \ldots ,$$

where X_0 is a selected "starting" value. Find the limitations, if any, on the starting value X_0 in order that the above iteration converges to the desired value $1/A$.

Problem 20.41 (Vietnam 1998). *Let $a \geq 1$ be a real number, and define the sequence $x_1, x_2, \ldots$ by $x_1 = a$ and*

$$x_{n+1} \;=\; 1 + \ln\left(\frac{x_n(x_n^2+3)}{3x_n^2+1}\right) .$$

Prove that this sequence has a finite limit, and determine it.

Problem 20.42 (CSM 1998). *Find all functions $f : \mathbb{N}^* \to \mathbb{N}^* \setminus \{1\}$ such that for all $n \in \mathbb{N}^*$,*

$$f(n) + f(n+1) \;=\; f(n+2)\, f(n+3) - 168 .$$

Problem 20.43 (SPCMO 2001). *Find all functions $f : \mathbb{Z} \to \mathbb{Z}$ such that*

$$f(x + y + f(y)) \;=\; f(x) + 2y$$

for all integers x, y.

Problem 20.44 (RMM 2008). *Prove that every bijective function $f : \mathbb{Z} \to \mathbb{Z}$ can be written in the way $f = u + v$, where $u, v : \mathbb{Z} \to \mathbb{Z}$ are bijective functions.*

Problem 20.45 (Iran 1997). *Find all functions $f : \mathbb{N}^* \to \mathbb{N} \setminus \{1\}$ such that*

$$f(n+1) + f(n+3) \;=\; f(n+5)\, f(n+7) - 1375 ,$$

for all $n \geq 0$.

Problem 20.46 (IMO 1989 Longlist). *For $n \in \mathbb{N}^*$, let $X = \{1, 2, \ldots, n\}$ and k an integer such that $n/2 \leq k \leq n$. Determine, with proof, the number of all functions $f : X \to X$ that satisfy the following conditions:*
(a) $f^2 = f$;
(b) the number of elements in the image of f is k;
(c) for each y in the image of f, the number of all points in X such that $f(x) = y$ is at most 2.

Problem 20.47 (IMO 1978). *Let $\{f(n)\}$ be a strictly increasing sequence of positive integers:*

$$0 < f(1) < f(2) < f(3) < \cdots .$$

Of the positive integers not belonging to the sequence, the n-th in order is of magnitude is $f^2(n) + 1$. Determine $f(240)$.

Problem 20.48 (IMO 1971 Shortlist). *Let $T_k = k - 1$ for $k = 1, 2, 3, 4$ and*

$$T_{2k-1} = T_{2k-2} + 2^{k-2} , \quad T_{2k} = T_{2k-5} + 2^k ,$$

for $k \geq 3$. Show that

$$\begin{aligned} 1 + T_{2k-1} &= \left\lfloor \frac{12}{7} 2^{n-1} \right\rfloor , \\ 1 + T_{2k} &= \left\lfloor \frac{17}{7} 2^{n-1} \right\rfloor , \end{aligned}$$

for all k.

Problem 20.49 (IMO 2010). *Determine all functions $g : \mathbb{N}^* \to \mathbb{N}^*$ such that*

$$(g(m) + n)\ (g(n) + m)$$

is a perfect square for all $m, n \in \mathbb{N}^*$.

Problem 20.50 (Bulgaria 2001). *A sequence $\{a_n\}$ is defined by*

$$a_{n+2} = 3a_{n+1} - 2a_n ,$$

for $n \geq 1$ and $a_1 = \lambda$, $a_2 = 5\lambda - 2$, where λ is a real number.

(a) Find all values of λ such that the sequence $\{a_n\}$ is convergent.

(b) Prove that if $\lambda = 1$ then

$$a_{n+2} = \left\lfloor \frac{7a_{n+1}^2 - 8a_n a_{n+1}}{1 + a_n + a_{n+1}} \right\rfloor .$$

Problem 20.51 ([2], Problem 890). *Let a be any positive integer and let H_n be the sequence recursively defined by $H_0 = 0$, $H_1 = 1$, and*

$$H_{n+2} = a H_{n+1} + H_n \quad \forall n \geq 0 .$$

(For example, for $a = 1$ these are the Fibonacci numbers and for $a = 2$ the Pell numbers.) Let k be a non-negative integer and let G_n be the sequence defined by

$$G_n = H_n H_{n+k} .$$

Find a positive integer m and constants $c_1, c_2, \ldots, c_m$ (depending only on a and k) such that

$$G_n = \sum_{i=1}^{m} c_i G_{n-i} ,$$

for all $n \geq m$ or prove that for all m such constants cannot exist.

Problem 20.52 (Netherlands 2009). *Find all functions $f : \mathbb{N}^* \to \mathbb{N}^*$ that satisfy the following two conditions:*

(a) $f(n)$ is a perfect square for all $n \in \mathbb{N}^*$;

(b) $f(m+n) = f(m) + f(n) + 2mn$, for all $m, n \in \mathbb{N}^*$.

Problem 20.53 (Netherlands TST 2009). *Find all functions $f : \mathbb{Z} \to \mathbb{Z}$ that satisfy the condition*

$$f(m+n) + f(mn-1) = f(m) f(n) + 2$$

for all $m, n \in \mathbb{Z}$.

Problem 20.54 (Putnam 1966). *Let $0 < x_1 < 1$ and $x_{n+1} = x_n(1 - x_n)$, $n = 1, 2, \cdots$. Show that*

$$\lim_{n\to\infty} n x_n = 1 .$$

Problem 20.55 (USA TST 2004). *Define the function $f : \mathbb{N} \to \mathbb{Q}$ by $f(0)$ and*

$$f(3n+k) = -\frac{3}{2} f(n) + k , \quad k = 0, 1, 2 , \quad n \in \mathbb{N} .$$

Prove that f is injective and find $f(\mathbb{N})$.

20.4 Functional Equations With Parameters

Problem 20.56 ([13], Pr 477). *Show that there is only one value of the constant b for which there exists a real function defined for all real numbers with the property that, for all real x and y,*

$$f(x-y) = f(x) - f(y) + b\,xy\,.$$

Problem 20.57. *Show that there are two continuous functions $f : \mathbb{R} \to \mathbb{R}$ which solve the functional relation*

$$f(x) + f(y) = f(x)\,f(y) - xy - 1\,, \forall x, y \in \mathbb{R}\,.$$

Problem 20.58. *Find all functions $f : \mathbb{R}^*_+ \to \mathbb{R}^*_+$ such that for all $x, y \in \mathbb{R}^*_+$,*

$$f(x+y) + f(x)\,f(y) = f(x) + f(y) + f(xy)\,.$$

Problem 20.59. *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that $\forall x, y \in \mathbb{R}$,*

$$f(x+y) + f(x)\,f(y) = f(xy) + 1\,.$$

Problem 20.60 (IMO 2005 Shortlist). *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that $\forall x, y \in \mathbb{R}$,*

$$f(x+y) + f(x)\,f(y) = f(xy) + 2xy + 1\,.$$

Problem 20.61. *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that $\forall x, y \in \mathbb{R}$,*

$$f(x+y) + f(x)\,f(y) = f(xy) + g(x, y)\,,$$

where g is a given function symmetric in x, y and $g(0,0) = 1$.

Problem 20.62. *Let $f, g_1, g_2, h_1, h_2 : \mathbb{R} \to \mathbb{R}$ be continuous functions that satisfy the functional relation*

$$f(x+y) \;=\; g_1(x)\, g_2(y) + h_1(x)\, h_2(y)\ ,$$

$\forall x, y$. Find all such functions.

Problem 20.63 (CSM 1999)**.** *Find all functions $f : (0, +\infty) \to \mathbb{R}$ such that*

$$f(x) - f(y) \;=\; (y-x)\, f(xy)\ ,$$

for all $x, y > 1$.

Problem 20.64. *Let S^1 be the 1-dimensional circle, that is the unit interval with end points identified. Find all continuous functions $\Phi : S^1 \to S^1$ such that*

$$\Phi(\theta_1 + \theta_2) \;=\; \Phi(\theta_1) + \Phi(\theta_2)\ ,$$

for all $\theta_1, \theta_2 \in S^1$.

Problem 20.65 (NMC 1998)**.** *Find all functions from the rational numbers to the rational numbers satisfying*

$$f(x+y) + f(x-y) \;=\; 2f(x) + 2f(y)\ ,$$

for all rational x and y.

Problem 20.66 (Iran 1997)**.** *Suppose $f : \mathbb{R}_+^* \to \mathbb{R}_+^*$ is a decreasing function such that for all $x, y \in \mathbb{R}_+^*$*

$$f(x+y) + f\Big(f(x) + f(y)\Big) \;=\; f\Bigg(f\Big(x + f(y)\Big) + f\Big(y + f(x)\Big)\Bigg)\,.$$

Prove that

$$f(x) \;=\; f^{-1}(x)\,.$$

Problem 20.67 (BMO 2007). *Find all real functions defined on* $\mathbb{R}$ *such that*

$$f(f(x)+y) \;=\; f(f(x)-y)+4f(x)\,y\;,$$

for all real numbers x, y.

Problem 20.68 (IMO 2004 Shortlist). *Find all functions* $f:\mathbb{R}\to\mathbb{R}$ *satisfying the equation*

$$f\left(x^2+y^2+2f(xy)\right) \;=\; f(x+y)^2$$

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

Problem 20.69 (IMO 1992). *Find all functions* $f:\mathbb{R}\to\mathbb{R}$ *such that*

$$f\left(x^2+f(y)\right) \;=\; y+f(x)^2$$

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

Problem 20.70 (IMO 1999). *Determine all functions* $f:\mathbb{R}\to\mathbb{R}$ *such that*

$$f(x-f(y)) \;=\; f(f(y))+xf(y)+f(x)-1$$

for all real numbers x, y.

Problem 20.71 (Iran 1999; BH TST 2008). *Find all functions* $f:\mathbb{R}\to\mathbb{R}$ *satisfying*

$$f(f(x)+y) \;=\; f(x^2-y)+4f(x)y\;,$$

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

Problem 20.72 (Italy 1999). *(a) Find all the strictly monotonic functions $f : \mathbb{R} \to \mathbb{R}$ such that*

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

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

(b) Prove that for every integer $n > 1$ there do not exist strictly monotonic functions $f : \mathbb{R} \to \mathbb{R}$ such that

$$f(x + f(y)) = f(x) + y^n,$$

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

Problem 20.73 (Vietnam 1999). *Let $f(x)$ be a continuous function defined on $[0, 1]$ such that $f(0) = f(1) = 0$ and*

$$2f(x) + f(y) = 3f\left(\frac{2x + y}{3}\right),$$

for all $x, y \in [0, 1]$. Prove that $f(x) = 0$ for all $x, y \in [0, 1]$.

Problem 20.74 (India 2000). *Suppose $f : \mathbb{Q} \to \{0, 1\}$ is a function with the property that for $x, y \in \mathbb{Q}$,*

$$\text{if } f(x) = f(y) \text{ then } f(x) = f\left(\frac{x + y}{2}\right) = f(y).$$

If $f(0) = 0$ and $f(1) = 1$ show that $f(q) = 1$ for all rational numbers q greater than or equal to 1.

Problem 20.75 (Korea 2000). *Find all functions $f : \mathbb{R} \to \mathbb{R}$ satisfying*

$$f(x^2 - y^2) = (x - y)(f(x) + f(y))$$

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

Problem 20.76 (APMO 2002). *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that there are only finitely many $s \in \mathbb{R}$ for which $f(s) = 0$ and*

$$f(x^4 + y) = x^3 f(x) + f^2(y)$$

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

Problem 20.77 (India 2005). *Find all functions $f:\mathbb{R}\to\mathbb{R}$ such that*

$$f(x^2 + y f(z)) = x f(x) + z f(y) ,$$

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

Problem 20.78 (Poland 1990). *Find all functions $f:\mathbb{R}\to\mathbb{R}$ such that*

$$(x-y) f(x+y) - (x+y) f(x-y) = 4xy (x^2 - y^2) ,$$

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

Problem 20.79 (Ukraine TST 2007). *Find all functions $f:\mathbb{Q}\to\mathbb{Q}$ such that*

$$f(x^2 + y + f(xy)) = 3 + (x + f(y) - 2) f(x)$$

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

Problem 20.80 (Ukraine TST 2008). *Find all functions $f:\mathbb{R}_+^*\to\mathbb{R}_+^*$ such that*

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

for all pairs of positive reals x and y.

Problem 20.81 (Japan 2004). *Find all functions $f:\mathbb{R}\to\mathbb{R}$ such that*

$$f(xf(x) + f(y)) = f(x)^2 + y$$

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

Problem 20.82 (Japan 2006). *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(x)^2 + 2yf(x) + f(y) = f(f(x) + y)$$

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

Problem 20.83 (Japan 2008). *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(x + y)\, f(f(x) - y) = x\, f(x) - y\, f(y)$$

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

Problem 20.84 (Japan 2009). *Find all functions $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that*

$$f(x^2) + f(y) = f(x^2 + y + xf(4y))$$

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

Problem 20.85 (IMO 1969 Longlist). *Find all functions f defined for all x that satisfy the condition*

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

for all x and y. Prove that exactly two of them are continuous.

Problem 20.86 (IMO 1976 Longlist). *Find the function $f(x)$ defined for all real values of x that satisfies the conditions*
(a) $f(x + 2) - f(x) = x^2 + 2x + 4$, for all x
(b) $f(x) = x^2$, if $x \in [0, 2)$.

Problem 20.87 (IMO 1979 Shortlist). *Prove that the functional equations*

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

and

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

are equivalent.

Problem 20.88 (Austria 2001)**.** *Determine all functions $f : \mathbb{R} \to \mathbb{R}$ such that for all real numbers x and y the functional equation*

$$f(f(x)^2 + f(y)) = x f(x) + y$$

is satisfied.

Problem 20.89 (IMO 1978 Longlist)**.** *(Version 1) Let $c, s : \mathbb{R}^* \to \mathbb{R}$ be non-constant in any interval and satisfy*

$$c\left(\frac{x}{y}\right) = c(x)\,c(y) - s(x)\,s(y)\,, \quad \forall x, y \in \mathbb{R}^*\,.$$

Prove that:
(a) $c(1/x) = c(x)$, $s(1/x) = -s(x)$, for any $x \neq 0$;
(b) $c(1) = 1$, $s(1) = s(-1) = 0$;
(c) c, s are either both even or both odd function.
(d) Find functions c and s that also satisfy $c(x) + s(x) = x^n$ for all $x \neq 0$, where n is a given positive integer.

(Version 2) Given a non-constant function $f : \mathbb{R}^*_+ \to \mathbb{R}$ such that $f(xy) = f(x)f(y)$ for any $x, y \neq 0$, find functions $c, s : \mathbb{R}^*_+ \to \mathbb{R}$ that satisfy

$$c\left(\frac{x}{y}\right) = c(x)\,c(y) - s(x)\,s(y)\,, \quad \forall x, y \in \mathbb{R}^*\,.$$

and

$$c(x) + s(x) = f(x)\,, \quad \forall x \in \mathbb{R}^*\,.$$

Problem 20.90 (IMO 1989 Longlist)**.** *Let $f : \mathbb{Q} \to \mathbb{R}_+$ satisfy the following conditions:*
(a) $f(0) = 0$;
(b) $f(ab) = f(a)\,f(b)$;
(c) $f(a + b) \leq f(a) + f(b)$;
(d) $f(m) \leq 1989$, $\forall m \in \mathbb{Z}$.

Prove that

$$f(a + b) = \max\{f(a),\ f(b)\}\,, \quad \textit{if } f(a) \neq f(b)\,.$$

Problem 20.91 (IMO 2002 Shortlist). *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(f(x) + y) = 2x + f(f(y) - x), \quad \forall x, y \in \mathbb{R}.$$

Problem 20.92 (IMO 2002). *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that*

$$[f(x) + f(z)]\,[f(y) + f(t)] = f(xy - zt) + f(xt + yz), \quad \forall x, y, z, t \in \mathbb{R}.$$

Problem 20.93 (IMO 2010). *Determine all functions $f : \mathbb{R} \to \mathbb{R}$ such that the equality*

$$f(\lfloor x \rfloor y) = f(x) \lfloor f(y) \rfloor$$

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

Problem 20.94 ([1], Problem 11345). *Find all increasing functions $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(x + f(y)) = f^2(x) + f(y),$$

for all real x and y.

Problem 20.95 (Ukraine 2001). *Does there exist a function $f : \mathbb{R} \to \mathbb{R}$ such that for all $x, y \in \mathbb{R}$ the following equality holds?*

$$f(xy) = \max\{f(x), y\} + \min\{f(y), x\}.$$

Problem 20.96 (BH 1997). *Let $f : A \to \mathbb{R}$, $A \subseteq \mathbb{R}$, be a function with the following characteristic:*

$$f(x + y) = f(x) f(y) - f(xy) + 1,$$

for all $x, y \in A$.

(a) If $f : A \to \mathbb{R}$, $\mathbb{N} \subset A \subseteq \mathbb{R}$, is such a function, prove that the following is true:

$$f(n) = \begin{cases} \frac{c^{n+1}-1}{c-1}, & n \in \mathbb{N}, \quad c \neq 1, \\ n + 1, & n \in \mathbb{N}, \quad c = 1, \end{cases}$$

with $c = f(1) - 1$.

(b) Solve the given functional equation for $A = \mathbb{N}$.

(c) If $A = \mathbb{Q}$, find all the functions f which satisfy the given equation and the condition $f(1997) \neq f(1998)$.

20.5 Functional Equations with No Parameters

Problem 20.97 ([3], Problem 2828). *Find all functions $f : \mathbb{R} \setminus \{1\} \to \mathbb{R}$ satisfying the functional equation*

$$f(x) + a\, f\left(\frac{x+b}{x-1}\right) \;=\; c - x\,,$$

where $a \neq \pm 1$, $b \neq -1$ constants.

Problem 20.98 (Turkey TST 1999). *Determine all functions $f : \mathbb{R} \to \mathbb{R}$ such that the set*

$$\left\{\frac{f(x)}{x} \;:\; x \neq 0 \quad \text{and} \quad x \in \mathbb{R}\right\}$$

is finite, and for all $x \in \mathbb{R}$

$$f(x - 1 - f(x)) \;=\; f(x) - x - 1\,.$$

Problem 20.99 (Switzerland 1999). *Determine all functions $f : \mathbb{R}^* \to \mathbb{R}$ satisfying*

$$\frac{1}{x}\, f(-x) + f\left(\frac{1}{x}\right) \;=\; x\,,$$

for all $x \in \mathbb{R}^*$.

Problem 20.100 (Putnam 1971). *Find all functions $f : \mathbb{R} \setminus \{0, 1\} \to \mathbb{R}$ satisfying the functional equation*

$$f(x) + f\left(\frac{x-1}{x}\right) \;=\; 1 + x\,.$$

Problem 20.101 (Austria 1985). *Determine all functions $f : \mathbb{R} \to \mathbb{R}$ satisfying the functional equation*

$$x^2 f(x) + f(1-x) \;=\; 2x - x^4\,, \quad \forall x \in \mathbb{R}\,.$$

Problem 20.102 (APMO 1989). *Determine all non-decreasing functions $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(x) + g(x) = 2x , \quad \forall x ,$$

and where g is defined by

$$f(g(x)) = g(f(x)) = x .$$

Problem 20.103. *Find all the solutions $f : I \to I$ of the functional equation*

$$f^3(x) = 0 .$$

Problem 20.104. *Find all the solutions of the functional equation*

$$f^3(x) = x .$$

Problem 20.105. *Is there a continuous solution of the functional equation*

$$f^4(x) = x$$

that is not a solution of $f^2(x) = x$? If no, why? If yes, give an example.

Problem 20.106 (Putnam 1996). *Let $c \geq 0$ be a constant. Give a complete description, with proof, of the set of all continuous functions $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(x) = f\left(x^2 + c\right) ,$$

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

Problem 20.107 (Putnam 2000). *Let $f(x)$ be a continuous function such that*

$$f\left(2x^2 - 1\right) = 2x\, f(x)$$

for all x. Show that $f(x) = 0$ for $-1 \leq x \leq 1$.

Problem 20.108 (Korea 1999). *Suppose that for any real x with $|x| \neq 1$, a function $f(x)$ satisfies*

$$f\left(\frac{x-3}{x+1}\right) + f\left(\frac{3+x}{1-x}\right) = x .$$

Find all possible $f(x)$.

Problem 20.109 (India 1992). *Determine all functions $f : \mathbb{R} \setminus [0,1] \to \mathbb{R}$ such that*

$$f(x) + f\left(\frac{1}{1-x}\right) = \frac{2(1-2x)}{x(1-x)} .$$

Problem 20.110 (Israel 1995). *For a given real number α, find all functions $f : (0,+\infty) \to (0,+\infty)$ satisfying*

$$\alpha x^2 f\left(\frac{1}{x}\right) + f(x) = \frac{x}{x+1} .$$

Problem 20.111 (Poland 1993; India 1994). *Find all real-valued functions f on the reals such that*
(a) $f(-x) = -f(x)$, for all x;
(b) $f(x+1) = f(x) + 1$, for all x;
(c) $f\left(\frac{1}{x}\right) = \frac{f(x)}{x^2}$, for $x \neq 0$.

Problem 20.112 (Poland 1992). *Find all functions $f : \mathbb{Q}_+^* \to \mathbb{Q}_+^*$ such that*

$$f(x+1) = f(x) + 1$$

and

$$f(x^3) = f(x)^3 ,$$

for all x.

Problem 20.113 (IMO 1979 Shortlist)**.** *For all rational x satisfying $0 \le x \le 1$, f is defined by*

$$f(x) = \begin{cases} \frac{f(2x)}{4}\,, & 0 \le x \le \frac{1}{2}\,, \\ \frac{3+f(2x-1)}{4}\,, & \frac{1}{2} \le x \le 1\,. \end{cases}$$

Given that $x = 0.b_1b_2b_3\ldots$ is the binary representation of x, find $f(x)$.

Problem 20.114 (Korea 1993)**.** *Let n be a given natural number. Find all continuous functions $f(x)$ satisfying*

$$\sum_{k=0}^{n} \binom{n}{k} f\left(x^{2^k}\right) = 0\,.$$

Problem 20.115 (Croatia 1995)**.** *Let $t \in (0,1)$ be a given number. Find all functions $f : \mathbb{R} \to \mathbb{R}$ continuous at $x = 0$ satisfying*

$$f(x) - 2f(tx) + f(t^2x) = x^2\,, \quad \forall x \in \mathbb{R}\,.$$

Problem 20.116 (IMO 1979 Longlist)**.** *Let E be the set of all bijective functions from $\mathbb{R}$ to $\mathbb{R}$ satisfying*

$$f(t) + f^{-1}(t) = 2t\,, \quad \forall t \in \mathbb{R}\,.$$

Find all elements of E that are monotonic functions.

20.6 Fixed Points and Cycles

Problem 20.117. *Let $f : [a,b] \to [a,b]$ be a monotonic function, not necessarily continuous. Prove that f has a fixed point.*

Problem 20.118. *Let $f : [a,b] \to [a,b]$ be a continuous surjective map. Prove that the function $f^2(x))$ has at least two fixed points.*

Problem 20.119 ([2], Problem 841). *Assume that the quadratic polynomial $f(x) = ax^2 + bx + c$, $a \neq 0$, has two fixed points x_1 and x_2, $x_1 \neq x_2$. If 1 and -1 are two fixed points of the function $f^2(x)$, but not of $f(x)$, then find the exact values of x_1 and x_2.*

Problem 20.120 ([4], Problem Q982). *It is well known that if $f : [0,1] \to [0,1]$ is a continuous function, then f must have at least one fixed point. However, the example $f(x) = x^2$, which only has 0 and 1 as fixed points shows that the set of fixed points need not be an interval.*

Let $f : [0,1] \to [0,1]$ be a function with $|f(x) - f(y)| \leq |x - y|$ for all$x, y \in [0,1]$. Prove that the set of all fixed points of f is either a single point or an interval.

Problem 20.121 (IMSUC 2007). *Let*

$$T = \{(tq, 1-t) \in \mathbb{R}^2 \mid t \in [0,1],\ q \in \mathbb{Q}\} .$$

Prove that each continuous function $f : T \to T$ has a fixed point.

Problem 20.122 (IMO 1976 Shortlist). *Let $I = (0,1]$. For a given number $a \in (0,1]$ we define the map $T : I \to I$ as follows:*

$$T(x) = \begin{cases} x + (1-a)\,, & \text{if } 0 < x \leq a\,, \\ x - a\,, & \text{if } a < x \leq 1\,. \end{cases}$$

Show that for every interval $J \subset I$ there exists an integer $n > 0$ such that $T^n(J) \cap J \neq \emptyset$.

Problem 20.123 (IMO 1976 Longlist). *Let $Q = [0,1] \times [0,1]$ and $T : Q \to Q$ be defined as follows:*

$$T(x,y) = \begin{cases} \left(2x, \frac{y}{2}\right)\,, & \text{if } 0 \leq x \leq 1/2\,, \\ \left(2x - 1, \frac{y+1}{2}\right)\,, & \text{if } 1/2 < x \leq 1\,. \end{cases}$$

Show that for every disc $D \subset Q$ there exists an integer $n > 0$ such that $T^n(D) \cap D \neq \emptyset$.

Problem 20.124 (IMO 1990 Shortlist)**.** *Let a, b be natural numbers with $1 \le a \le b$ and $M = \left\lfloor \frac{a+b}{2} \right\rfloor$. Define the function $f : \mathbb{Z} \to \mathbb{Z}$ by*

$$f(n) = \begin{cases} n + a\,, & \text{if } n < M\,, \\ n - b\,, & \text{if } n \ge M\,. \end{cases}$$

Find the smallest natural number $k > 0$ such that $f^k(0) = 0$.

Problem 20.125 ([4], Problem 2003)**.** *Let $(X, \langle\ \rangle)$ be a real inner product space, and let*

$$B = \{x \in X : \|x\| \le 1\}$$

be the unit ball in X, where $\|x\| = \sqrt{\langle x, x\rangle}$. Let $f : B \to B$ be a function satisfying

$$\|f(x) - f(y)\| \le \|x - y\|\,, \quad \forall x, y \in B\,.$$

Prove that the set of fixed points of f is convex.

Problem 20.126 ([1], Problem E3342)**.** *Let $S = \{1, 2, \ldots, n^2\}$ and define a permutation $f : S \to S$ as follows. Take n^2 cards numbered from 1 to n^2 and lay them in a square array, with the i-th row containing cards $(i-1)n+1, (i-1)n+2, \ldots,$ in in order from left to right. Pick up the cards along rising diagonals, starting with the upper left-hand corner. If card j is the k-th card picked up, put $f(j) = k$. For example, if $n = 3$, then $f(1) = 1$, $f(4) = 2$, $f(2) = 3$, $f(7) = 4$, $f(5) = 5$, $f(3) = 6$, $f(8) = 7$, $f(6) = 8$, $f(9) = 9$.*

For each value of n both 1 and n^2 are fixed points points of f. If n is odd, then $(n^2+1)/2$ is also a fixed point of f. Characterize those n for which f has a fixed point not in $\{1,\ n^2,\ (n^2+1)/2\}$.

Problem 20.127 ([1], Problem E3428)**.** *Let I be a non-empty interval on the real line. Let $f : I \to I$ be a continuous function having the property that for each $x \in I$ there exists a positive integer $n = n(x)$ with $f^n(x) = x$. For given I characterize all such functions.*

Problem 20.128 (Sweden 1993)**.** *Let a and b be real numbers and let*

$$f(x) = \frac{1}{a\,x + b}\,.$$

For which a and b are there three distinct real numbers x_1, x_2, x_3 such that

$$f(x_1) = x_2\,, \quad f(x_2) = x_3\,, \quad f(x_3) = x_1\ ?$$

20.7 Existence of Solutions

Problem 20.129 (Bulgaria 1998). *Prove that there does not exist a function $f : \mathbb{R}_+^* \to \mathbb{R}_+^*$ such that*

$$f(x)^2 \geq f(x+y)\,(f(x)+y)\,,$$

for any $x, y \in \mathbb{R}_+^*$.

Problem 20.130 (Iran 1998). *Let $f_1, f_2, f_3 : \mathbb{R} \to \mathbb{R}$ be functions such that $a_1 f_1 + a_2 f_2 + a_3 f_3$ is monotonic for all $a_1, a_2, a_3 \in \mathbb{R}$. Prove that there exist $c_1, c_2, c_3 \in \mathbb{R}$, not all zero, such that*

$$c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) \;=\; 0\,,$$

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

Problem 20.131 (Belarus 2000). *Does there exist a function $f : \mathbb{N}^* \to \mathbb{N}^*$ such that*

$$f(f(n-1)) \;=\; f(n+1) - f(n)$$

for all $n \geq 2$?

Problem 20.132 (LMO 1991). *Does there exist a function $F : \mathbb{N}^* \to \mathbb{N}^*$ such that for any natural number x,*

$$F(F(F(\cdots F(x)\cdots))) \;=\; x+1\ ?$$

Here F is applied F times.

Problem 20.133 (IMO 1989 Shortlist). *Let $g : \mathbb{C} \to \mathbb{C}$, $\omega \in \mathbb{C}$, $a \in \mathbb{C}$, $\omega^3 = 1$, and $\omega \neq 1$. Show that there is one and only one function $f : \mathbb{C} \to \mathbb{C}$ such that*

$$f(z) + f(\omega z + a) \;=\; g(z)\,, \quad z \in \mathbb{C}\,.$$

Problem 20.134 (IMO 1967 Longlist). *The function $\phi(x, y, z)$ defined for all triples (x, y, z) of real numbers is such that there are two functions f and g defined for all pairs of real numbers such that*

$$\phi(x, y, z) = f(x + y, z) = g(x, y + z) ,$$

for all real numbers x, y, z. Show that there is a function h of one real variable such that

$$\phi(x, y, z) = h(x + y + z) ,$$

for all real numbers x, y, z.

Problem 20.135 (IMO 1992 Shortlist). *Let $\mathbb{R}_+$ be the set of all non-negative real numbers. Given two positive real numbers a and b, suppose that a mapping $f : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies the functional equation*

$$f(f(x)) + a f(x) = b(a + b)x .$$

Prove that there exists a unique solution of this equation.

Problem 20.136 (IMO 1993). *Let $\mathbb{N}^* = \{1, 2, 3, \ldots\}$. Determine whether or not there exists a function $f : \mathbb{N}^* \to \mathbb{N}^*$ such that $f(1) = 2$ and*

$$\begin{aligned} f(f(n)) &= f(n) + n , \quad \text{for all } n \text{ in } \mathbb{N}^* , \\ f(n) &< f(n + 1) , \quad \text{for all } n \text{ in } \mathbb{N}^* . \end{aligned}$$

Problem 20.137 (Russia 2000). *Is there a function $f : \mathbb{R} \to \mathbb{R}$ such that*

$$|f(x + y) + \sin x + \sin y| < 2$$

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

Problem 20.138 (Turkey 2000). *Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that*

$$|f(x + y) - f(x) - f(y)| \leq 1$$

for all $x, y \in \mathbb{R}$. Show that there exists a function $g : \mathbb{R} \to \mathbb{R}$ with

$$|f(x) - g(x)| \leq 1$$

for all $x \in \mathbb{R}$ and

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

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

Problem 20.139 (Romania 2000). *Prove that there is no function $f : (0, +\infty) \to (0, +\infty)$ such that*

$$f(x + y) \geq f(x) + y f^2(x)$$

for all $x, y \in (0, +\infty)$.

Problem 20.140 (Romania 2005). *Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function.*
(a) Prove that f is continuous.
(b) Prove that there exists a unique function $g : [0, +\infty) \to \mathbb{R}$ such that for all $x \geq 0$

$$f(x + g(x)) \;=\; f(g(x)) - g(x) \;.$$

Problem 20.141 ([30], Problem 54). *Is there a non-trivial function $f(x)$, continuous on the whole line, which satisfies the functional equation*

$$f(x) + f(2x) + f(3x) \;=\; 0 \;,$$

for all x?

Problem 20.142 (Iran 1995). *Does there exist a function $f : \mathbb{R} \to \mathbb{R}$ that fulfils all of the following conditions:*

(a) $f(1) = 1$,
(b) there exists $M > 0$ such that $-M < f(x) < M$,
(c) if $x \neq 0$ then

$$f\left(x + \frac{1}{x^2}\right) \;=\; f(x) + f\left(\frac{1}{x}\right)^2 \;?$$

Problem 20.143 (IMO 1970 Longlist). *Let E be a finite set and $f : \mathcal{P}(E) \to \mathbb{R}_+$ such that for any two disjoint sets A, B of E,*

$$f(A \cup B) \;=\; f(A) + f(B) \;.$$

Prove that there exists a subset F of E such that if with each $A \subset E$ we associate a subset A' consisting of elements of A that are not in F, then $f(A) = f(A')$, and $f(A)$ is zero iff A is a subset of F.

Problem 20.144 ([2], Problem 869). *Prove that there does not exist a positive twice differentiable function f defined on $[0, +\infty)$ such that*

$$f(x)\, f''(x) \leq -1 \;, \quad \forall x \geq 0 \;.$$

20.8 Systems of Functional Equations

Problem 20.145 (APMC 1999). *Let $n \geq 2$ be a given integer. Determine all systems of n functions $(f_1, \ldots, f_n)$ where $f_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, \ldots, n$, such that for all $x, y \in \mathbb{R}$ the following equalities hold:*

$$\begin{aligned} f_1(x) - f_2(x) f_2(y) + f_1(y) &= 0\,, \\ f_2(x^2) - f_3(x) f_3(y) + f_2(y^2) &= 0\,, \\ \cdots \quad & \cdots \\ f_{n-1}(x^{n-1}) - f_n(x) f_n(y) + f_{n-1}(y^{n-1}) &= 0\,, \\ f_n(x^n) - f_1(x) f_1(y) + f_n(y^n) &= 0\,. \end{aligned}$$

Problem 20.146 (IMO 1977 Longlist). *Let $f : \mathbb{Q}^* \times \mathbb{Q}^* \to \mathbb{R}_+^*$ be a function satisfying the equations*

$$\begin{aligned} f(x_1 x_2, y) &= f(x_1, y)\, f(x_2, y)\,, \quad \forall x_1, x_2, y \\ f(x, y_1 y_2) &= f(x, y_1)\, f(x, y_2)\,, \quad \forall x, y_1, y_2 \\ f(x, 1-x) &= 1\,, \quad \forall x\,. \end{aligned}$$

Prove that

$$\begin{aligned} f(x, x) = f(x, -x) &= 1\,, \quad \forall x\,, \\ f(x, y)\, f(y, x) &= 1\,, \quad \forall x, y\,. \end{aligned}$$

Problem 20.147 (BMO 2003). *Find all functions $f : \mathbb{Q} \to \mathbb{R}$ which fulfill the following conditions:*

(a) $f(1) + 1 > 0$;

(b) $f(x + y) - x f(y) - y f(x) = f(x) f(y) - x - y + xy$, *for all $x, y \in \mathbb{Q}$;*

(c) $f(x) = 2f(x + 1) + x + 2$, *for every $x \in \mathbb{Q}$.*

Problem 20.148 (IMO 1989 Longlist). *Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(1) = 1$ and*

$$\begin{aligned} f(x + y) &= f(x) + f(y)\,, \quad \forall x, y\,, \\ f(x)\, f\left(\frac{1}{x}\right) &= 1\,, \quad \forall x \neq 0\,. \end{aligned}$$

Find f.

Problem 20.149 (Putnam 1977)**.** *Let $f, g, u : \mathbb{R} \to \mathbb{R}$ such that*

$$\begin{aligned} \frac{u(x+1)+u(x-1)}{2} &= f(x) , \\ \frac{u(x+4)+u(x-4)}{2} &= g(x) . \end{aligned}$$

Determine u in terms of f and g.

Problem 20.150 ([1], Problem 11053)**.** *Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that*

$$\begin{aligned} f\big(x + f(x)f(y)\big) &= f(x) + x\,f(y) , \\ f\big(x + f(xy)\big) &= f(x) + x\,f(y) , \end{aligned}$$

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

20.9 Conditional Functional Equations

Problem 20.151 (IMO 2008)**.** *Find all functions $f : (0, +\infty) \to (0, +\infty)$ such that*

$$\frac{f(w)^2 + f(x)^2}{f(y^2) + f(z^2)} = \frac{w^2 + x^2}{y^2 + z^2} ,$$

for all positive real numbers w, x, y, z satisfying $wx = yz$.

Problem 20.152 (Poland 1991)**.** *On the Cartesian plane consider the set V of all vectors with integer coordinates. Determine all functions $f : V \to \mathbb{R}$ satisfying the conditions:*
(a) $f(\vec{u}) = 1$ for each of the four vectors $\vec{u} \in V$ of unit length.
(b) $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$ for every two perpendicular vectors $\vec{x}, \vec{y} \in V$.
(The zero vector is considered to be perpendicular to every vector).

Problem 20.153 (IMO 1989 Longlist)**.** *Let $a \in (0, 1)$ and f a continuous function on $[0, 1]$ satisfying $f(0) = 0$, $f(1) = 1$ and*

$$f\left(\frac{x+y}{2}\right) = (1-a)\,f(x) + a\,f(y) , \quad \forall\, x \le y , \quad x, y \in [0, 1] .$$

Determine $f(1/7)$.

Problem 20.154 (IMO 2004). *Find all polynomials $P(x)$ with real coefficients that satisfy the equality*

$$P(a-b)+P(b-c)+P(c-a) \;=\; 2\,P(a+b+c)\,,$$

for all triples a, b, c of real numbers such that $ab+bc+ca=0$.

20.10 Polynomials

Problem 20.155 (IMO 1975). *Find all polynomials P, in two variables, with the following properties:*

(a) For a positive integer n and all real t, x, y

$$P(tx, ty) \;=\; t^n\,P(x, y)\,,$$

that is, P is homogeneous of degree n.

(b) For all real a, b, c,

$$P(b+c, a)+P(c+a, b)+P(a+b, c) \;=\; 0\,.$$

(c) $P(1,0) \;=\; 1$.

Problem 20.156 (Vietnam 1998). *Prove that for each positive odd integer n, there is exactly one polynomial $P(x)$ of degree n with real coefficients satisfying*

$$P\left(x-\frac{1}{x}\right) \;=\; x^n-\frac{1}{x^n}\,,\quad \forall x \in \mathbb{R}^*\,.$$

Determine if the above assertion holds for positive even integers n.

Problem 20.157 (SPCMO 1998). *Find all polynomials $P(x, y)$ in two variables such that for any x and y,*

$$P(x+y, y-x) \;=\; P(x-y)\,.$$

Problem 20.158 (Greece 2005). *Find the polynomial $P(x)$ with real coefficients such that $P(2)=12$ and*

$$P\left(x^2\right) \;=\; x^2\left(x^2+1\right)\,P(x)\,,$$

for each $x \in \mathbb{R}$.

Problem 20.159 (IMO 1992 Shortlist). *Let f, g and a be polynomials with real coefficients, f and g in one variable and a in two variables. Suppose*

$$f(x) - f(y) = a(x, y)\Big(g(x) - g(y)\Big),$$

for all $x, y \in \mathbb{R}$. Prove that there exists a polynomial h with

$$f(x) = h(g(x)), \quad \forall x \in \mathbb{R}.$$

Problem 20.160 (China 1999). *Let a be a real number. Let $\{f_n(x)\}$ be a sequence of polynomials such that*
(a) $f_0(x) = 1$ and
(b) $f_{n+1}(x) = x f_n(x) + f_n(ax)$, for $n = 0, 1, 2, \cdots$.
Prove that

$$f_n(x) = x^n f_n\left(\frac{1}{x}\right), \quad n = 0, 1, 2, \ldots,$$

and find an explicit expression for $f_n(x)$.

Problem 20.161 (Georgia TST 2005). *Find all polynomials with real coefficients for which the equality*

$$P(2P(x)) = 2P^2(x) + 2P(x)^2$$

holds for any real number x.

Problem 20.162 (Japan 1995). *Find all non-constant rational functions $f(x)$ of x satisfying*

$$f(x)^2 - a = f(x^2),$$

where a is a real number.

*A **rational function** of x is one that can be expressed as the ratio of two polynomials of x.*

Problem 20.163 (CSM 1999). *Find all positive integers k for which the following assertion holds: If $F(x)$ is a polynomial with integer coefficients which satisfies*

$$F(c) \leq k\,, \quad \textit{for all } c \in \{0, 1, \ldots, k+1\},$$

then

$$F(0) \;=\; F(1) \;=\; \cdots \;=\; F(k+1)\,.$$

Problem 20.164 (Bulgaria 1998). *The polynomials $P_n(x, y)$ for $n = 1, 2, \ldots$ are defined by*

$$\begin{aligned} P_1(x,y) &= 1\,, \\ P_{n+1}(x,y) &= (x+y-1)\,(y+1)\,P_n(x,y+2) + (y-y^2)\,P_n(x,y)\,. \end{aligned}$$

Prove that

$$P_n(x,y) \;=\; P_n(y,x)\,,$$

for all n and all x, y.

Problem 20.165 (CSM 1998). *A polynomial $P(x)$ of degree $n \geq 5$ with integer coefficients and n distinct integer roots is given. Find all integer roots of $P(P(x))$ given that 0 is a root of $P(x)$.*

Problem 20.166 (IMO 1976 Longlist). *Prove that if for a polynomial $P(x, y)$ we have*

$$P(x-1, y-2x+1) \;=\; P(x,y)\,,$$

then there exists a polynomial $Q(x)$ such that $P(x, y) = Q(y - x^2)$.

Problem 20.167 (IMO 1979 Shortlist). *Find all polynomials $P(x)$ with real coefficients for which*

$$P(x)\,P(2x^2) \;=\; P(2x^3 + x)\,.$$

Problem 20.168 (IMO 1978 Longlist)**.** *Let the polynomials*

$$P(x) \;=\; x^n + \sum_{k=0}^{n-1} a_k\, x^k\;, \quad P(x) \;=\; x^m + \sum_{k=0}^{m-1} b_k\, x^k\;,$$

be given satisfying the identity

$$P(x)^2 \;=\; (x^2-1)\, Q(x)^2 + 1\;.$$

Prove that

$$P'(x) \;=\; n\, Q(x)\;.$$

Problem 20.169 (IMO 1989 Longlist)**.** *Let $P(x)$ be o polynomial such that the following inequalities are satisfied:*

$$\begin{aligned} P(1) &> P(0) > 0\;, \\ P(2) &> 2P(1) - P(0)\;, \\ P(3) &> 3P(2) - 3P(1) + P(0)\;, \end{aligned}$$

and also

$$P(n+4) > 4P(n+3) - 6P(n+2) + 4P(n+1) - P(n)\;, \quad \forall n \in \mathbb{N}\;.$$

Prove that

$$P(n) > 0\;, \quad n \in \mathbb{N}\;.$$

Problem 20.170 (Putnam 1972)**.** *Let n be an integer greater than 1. Show that there exists a polynomial $P(x, y, z)$ with integral coefficients such that*

$$P(x^n, x^{n+1}, x^{n+2} + x) \;=\; x\;.$$

Problem 20.171 (USA 1984)**.** *$P(x)$ is a polynomial of degree $3n$ such that*

$$\begin{aligned} &P(0) \;=\; P(3) \;=\; \ldots \;=\; P(3n) \;=\; 2\;, \\ &P(1) \;=\; P(4) \;=\; \ldots \;=\; P(3n-2) \;=\; 1\;, \\ &P(2) \;=\; P(5) \;=\; \ldots \;=\; P(3n-1) \;=\; 0\;, \\ &P(3n+1) \;=\; 730\;. \end{aligned}$$

Determine n.

Problem 20.172 (IMO 1993). *Let $P(x) = x^n + 5x^{n-1} + 3$ where $n > 1$ is an integer. Prove that $P(x)$ cannot be expressed as the product of two polynomials, each of which has all its coefficients integers and degree at least 1.*

Problem 20.173 ([4], Problem 1807). *Let P be a polynomial with integer coefficients and let s be an integer such that for some positive integer n, the number $s^{n+1}\,P(s)^n$ is a positive zero of P. Prove that $P(2) = 0$.*

Problem 20.174 ([2], Problem 879). *Consider the polynomial*

$$f(x) \;=\; x^4 - 4ax^3 + 6b^2x^2 - 4c^3x + d^4\;,$$

where a, b, c, and d are positive real numbers. Prove that if f has four positive distinct roots, then $a > b > c > d$.

Problem 20.175 (Putnam 1976). *Let*

$$P(x,y) \;=\; x^2y + xy^2\;, \quad Q(x,y) \;=\; x^2 + xy + y^2\;.$$

For $n = 2, 3, \cdots$, let

$$\begin{aligned} F_n(x,y) &= (x+y)^n - (x^n + y^n)\;, \\ G_n(x,y) &= (x+y)^n + x^n + y^n\;. \end{aligned}$$

Prove that, for each n, either F_n or G_n is expressible as a polynomial in P and Q with integer coefficients.

Problem 20.176 (Putnam 1981). *Let $P(x)$ be a polynomial with real coefficients and form the polynomial*

$$Q(x) \;=\; (x^2+1)P(x)\,P'(x) + x\left[P(x)^2 + P'(x)^2\right]\;.$$

Given that $P(x)$ has n distinct roots real roots exceeding 1, prove or disprove that $Q(x)$ has at least $2n-1$ distinct real roots.

Problem 20.177 (HIMC 1999). *Let $P(x)$ be a polynomial whose degree is at least 2. Define the sequence $\{Q_n(x)\}$ by $Q_1(x) = P(x)$ and*

$$Q_{k+1}(x) \;=\; P(Q_k(x))\;, \quad k = 1, 2, \cdots\;.$$

Let r_n be the average of the roots of $Q_n(x)$. It is given that $r_{19} = 99$. Find r_{99}.

20.11 Functional Inequalities

Problem 20.178 (**Popoviciu's Theorem** [34]). *Let f be a convex real-valued function defined on an interval I. Then*

$$f(x)+f(y)+f(z)+3f\left(\frac{x+y+z}{3}\right)\geq 2\left[f\left(\frac{x+y}{2}\right)+f\left(\frac{y+z}{2}\right)+f\left(\frac{z+x}{2}\right)\right],$$

for all $x,y,z\in I$.

Problem 20.179 ([55]). *Let $f(x)$ be subadditive on $\mathbb{R}$ and let $f'(x)$ exist on $(0,+\infty)$. if $f(x)+f(-x)\leq 0$ for any $x\in(0,+\infty)$, then $f'(x)$ is decreasing on $(0,+\infty)$.*

Problem 20.180 (IMO 1977). *Let $f(n)$ be a function defined on the set of all positive integers and having all its values in the same set. prove that if*

$$f(n+1)>f(f(n))$$

for each positive integer n, then

$$f(n)\ =\ n\,,\quad \text{for each } n\,.$$

Problem 20.181 ([10], p. 35). *Determine $f:\mathbb{N}^*\to\mathbb{N}^*$ such that for all k,m,n,*

$$f(km)+f(kn)-f(k)\,f(mn)\geq 1\,.$$

Problem 20.182 (APMO 1994). *Let $f:\mathbb{R}\to\mathbb{R}$ such that*
(a) $f(x)+f(y)+1\geq f(x+y)\geq f(x)+f(y)$, $\forall x,y$;
(b) $f(0)\geq f(x)$, $\forall x\in[0,1)$;
(c) $f(1)=f(-1)=1$.
Find all such functions f.

Problem 20.183 (HK 1999). *Let f be a function defined on the reals with the following properties*
(A) $f(1) = 1$;
(B) $f(x+1) = x\,f(x)$;
(C) $f(x) = 10^{g(x)}$;
where $g(x)$ is a function defined on the reals satisfying

$$g(ty + (1-t)z) \leq t\,g(y) + (1-t)\,g(z)$$

for all real y, z and some $0 \leq t \leq 1$.
(a) Prove that

$$t\,[g(n) - g(n-1)] \leq g(n+t) - g(n) \leq t\,[g(n+1) - g(n)]$$

where n is an integer and $0 \leq t \leq 1$.
(b) Prove that

$$\frac{4}{3} \leq f\left(\frac{1}{2}\right) \leq \frac{4\sqrt{2}}{3}\,.$$

Problem 20.184 (IMSUC 2008). *Let $\gamma : [0,1] \to [0,1] \times [0,1]$ be a mapping such that for each $s, t \in [0,1]$*

$$|\gamma(s) - \gamma(t)| \leq M\,|s-t|^{\alpha}$$

in which α, M are fixed numbers. Prove that if γ is surjective, then $\alpha < 1/2$.

Problem 20.185 (IMO 2007 Shortlist; Ukraine TST 2008). *Consider those functions $f : \mathbb{N}^* \to \mathbb{N}^*$ which satisfy the condition*

$$f(m+n) \geq f(m) + f(f(n)) - 1$$

for all $m, n \in \mathbb{N}^*$. Find all possible values of $f(2007)$.

Problem 20.186 (Romania 1981). *Determine whether there exists an injective function $f : \mathbb{R} \to \mathbb{R}$ with the property that for all x,*

$$f(x^2) - f(x)^2 \geq \frac{1}{4}\,.$$

Problem 20.187 ([10], p. 227). *If f is a surjective function and g is a bijective function, both defined on $\mathbb{N}$ and satisfying the property*

$$f(n) \geq g(n)\,, \quad \forall n\,,$$

then $f = g$.

Problem 20.188 (Russia 2000). *Find all functions $f : \mathbb{R} \to \mathbb{R}$ that satisfy the inequality*

$$f(x+y) + f(y+z) + f(z+x) \geq 3\, f(x+2y+3z)$$

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

Problem 20.189 (Japan 2007). *Find all functions $f : \mathbb{R}_+^* \to \mathbb{R}$ such that*

$$\begin{aligned} f(x) + f(y) &\leq \frac{f(x+y)}{2}\,, \\ \frac{f(x)}{x} + \frac{f(y)}{y} &\geq \frac{f(x+y)}{x+y}\,, \end{aligned}$$

for all $x, y > 0$.

Problem 20.190 (IMO 1979 Longlist). *If a function satisfies the conditions*

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

and

$$f(x+y) \leq f(x) + f(y)\,, \quad \forall x, y \in \mathbb{R}\,,$$

show that

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

20.12 Functional Equations Containing Derivatives

Problem 20.191 (Romania 1999). *The function $f : \mathbb{R} \to \mathbb{R}$ is differentiable and*

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

for every number x. Prove that f is a linear function.

Problem 20.192 (Putnam 1941). *Show that any solution $f(t)$ of the functional equation*

$$f(x+y)\,f(x-y) = f(x)^2 + f(y)^2 - 1, \quad \forall x, y \in \mathbb{R},$$

is such that

$$f''(t) = \pm\, m^2 f(t),$$

(m a non-negative constant) assuming the existence and continuity of the second derivative. Deduce that $f(t)$ is one of the functions

$$\pm \cos(mt), \quad \pm \cosh(mt).$$

Problem 20.193 (Putnam 1997). *Let f be a twice-differentiable real-valued function satisfying*

$$f(x) + f''(x) = -\, x\, g(x)\, f'(x),$$

where $g(x) \geq 0$ for all real x. Prove that $|f(x)|$ is bounded.

Problem 20.194 (Poland 1977). *A function $h : \mathbb{R} \to \mathbb{R}$ is differentiable and satisfies $h(ax) = bh(x)$ for all x, where a and b are given positive numbers and $|a| \neq 0, 1$. Suppose that $h'(0) \neq 0$ and the function h' is continuous at $x = 0$. Prove that $a = b$ and that there is a real number c such that $h(x) = cx$ for all x.*

Problem 20.195 ([1], Problem E3338). *Let s and t be given real numbers. Find all differentiable functions f on the real line that satisfy*

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

for all real x, y with $x \neq y$.

20.13 Functional Relations Containing Integrals

Problem 20.196 ([5], Problem 71-1). *Prove that there is only one non-negative function F for which*

$$F\left\{\int_0^x F(u)du\right\} = x\,, \quad x \geq 0\,,$$

namely $F(x) = Ax^n$ for appropriate values of A and n.

Problem 20.197 (Romania 1999). *Let $\mathbb{R} \to \mathbb{R}$ be a monotonic function for which there exist $a, b, c, d \in \mathbb{R}$ with $a \neq 0$, $c \neq 0$ such that for all x the following equalities hold:*

$$\int_x^{x+\sqrt{3}} f(t)\,dt = a\,x + b\,,$$
$$\int_x^{x+\sqrt{2}} f(t)\,dt = c\,x + b\,.$$

Prove that f is a linear function.

Problem 20.198 (Putnam 1958). *Prove that if $f(x)$ is continuous for $a \leq x \leq b$ and*

$$\int_a^b x^n f(x)\,dx = 0\,, \quad n = 0, 1, 2, \ldots\,,$$

then $f(x)$ is identically zero on $a \leq x \leq b$.

Problem 20.199 (Romania 2004). *Find all continuous functions $f : \mathbb{R} \to \mathbb{R}$ such that for all $x \in \mathbb{R}$ and for all $n \in \mathbb{N}^*$ we have*

$$n^2 \int_x^{x+\frac{1}{n}} f(t)\,dt = n\,f(x) + \frac{1}{2}\,.$$

Problem 20.200 (Romania 2004). *Let $f : [0,1] \to \mathbb{R}$ be an integrable function such that*

$$\int_0^1 f(x)\,dx \;=\; \int_0^1 x\,f(x)\,dx \;=\; 1\,.$$

Prove that

$$\int_0^1 f(x)^2\,dx \geq 4\,.$$

Problem 20.201 (Putnam 1972). *Let $f(x)$ be an integrable function in $0 \leq x \leq 1$ and suppose*

$$\int_0^1 f(x)\,x^k\,dx \;=\; 0\,, \quad k = 0, 1, \ldots, n-1\,,$$

and

$$\int_0^1 f(x)\,x^n\,dx \;=\; 1\,.$$

Show that

$$|f(x)| \geq 2^n\,(n+1)\,,$$

in a set of positive measure.

Problem 20.202 (Putnam 1973). *(a) On [0,1], let f have a continuous derivative satisfying $0 < f'(x) \leq 1$. Also suppose $f(0) = 0$. prove that*

$$\left(\int_0^1 f(x)\,dx\right)^2 \geq \int_0^1 f(x)^3\,dx\,.$$

(b) Show an example in which equality occurs.

Problem 20.203 ([4], Problem 1824). *Let f be a continuous real-valued function defined on [0, 1] and satisfying*

$$\int_0^1 f(x)\,dx \;=\; \int_0^1 x f(x)\,dx\,.$$

Prove that there exists a real number c, $0 < c < 1$, such that

$$c f(c) \;=\; \int_0^c x f(x)\,dx\,.$$

Problem 20.204 ([2], Problem 925). *Let f be a twice differentiable function on $\mathbb{R}$ with f continuous on [0,1] such that*

$$\int_0^1 f(x)\,dx \;=\; 2\int_{1/4}^{3/4} f(x)\,dx\ .$$

Prove that there exists an $x_0 \in (0,1)$ such that $f''(x_0) = 0$.

Problem 20.205 ([2], Problem 898). *Suppose f is a function defined on an open interval I such that $f''(x) \geq 0$ for all $x \in I$, and $[a,b] \subset I$. Prove that*

$$\int_0^1 f(a+(b-a)y)\,dy \geq \int_0^1 f\left(\frac{3a+b}{4}+\frac{b-a}{2}y\right)dy\ .$$

Part VII

AUXILIARY MATERIAL

ACRONYMS & ABBREVIATIONS

APMO	Asian Pacific Mathematical Olympiad
APMC	Austrian-Polish Mathematics Competition
BH	Bosnia Herzegovina
BMO	Balkan Mathematical Olympiad
BWMC	Baltic Way Mathematical Contest
CSM	Czech and Slovak Match
HIMC	Hungary-Israel Mathematical Competition
HK	Hong Kong
IMSUC	Iranian Mathematical Society Undergraduate Competition
IMO	International Mathematical Olympiad
LMO	Leningrad Mathematical Olympiad
NMC	Nordic Mathematical Contest
Putnam	W.L. Putnam Mathematical Competition
RMM	Romanian Masters in Mathematics
ROM	Republic of Moldova
SPCMO	St Petersburg City Mathematical Olympiad
TST	Team Selection Test
UK	United Kingdom

When the name of a country is given, it is implied that the problem was given in the national competition of that country. However, if the competition separates the exams for different grades, no effort has been made to include the corresponding grade.

SET CONVENTIONS

$\mathbb{N}$ the set of natural numbers including 0

$\mathbb{N}^*$ the set of positive natural numbers

In other books the symbol $\mathbb{N}$ may stand for the positive natural numbers, while $\mathbb{N}_0$ may be used to indicate the set of natural numbers including 0.

$\mathbb{Z}$ the set of integer numbers

$\mathbb{Z}^*$ the set of non-zero integer numbers

$\mathbb{Z}_+$ the set of non-negative integer numbers

$\mathbb{Z}_-$ the set of non-positive integer numbers

$\mathbb{Z}_+^*$ the set of positive natural numbers

$\mathbb{Z}_-^*$ the set of negative integer numbers

In other books the symbols $\mathbb{Z}_+$, $\mathbb{Z}_-$ may stand for the positive, negative integer numbers respectively.

$\mathbb{Q}$ the set of rational numbers

$\mathbb{Q}^*$ the set of non-zero rational numbers

$\mathbb{Q}_+$ the set of non-negative rational numbers

$\mathbb{Q}_-$ the set of non-positive rational numbers

$\mathbb{Q}_+^*$ the set of positive rational numbers

$\mathbb{Q}_-^*$ the set of negative rational numbers

In other books the symbols $\mathbb{Q}_+$, $\mathbb{Q}_-$ may stand for the positive, negative rational numbers respectively.

$\mathbb{R}$ the set of real numbers

$\mathbb{R}^*$ the set of non-zero real numbers

$\mathbb{R}_+$ the set of non-negative real numbers

$\mathbb{R}_-$ the set of non-positive real numbers

$\mathbb{R}_+^*$ the set of positive real numbers

$\mathbb{R}_-^*$ the set of negative real numbers

$\overline{\mathbb{R}}$ the set of extended real numbers, i.e. $\mathbb{R}$ with $\pm\infty$ included

In other books the symbols $\mathbb{R}_+$, $\mathbb{R}_-$ may stand for the positive, negative real numbers respectively, while the symbols $\overline{\mathbb{R}}_+$, $\overline{\mathbb{R}}_-$ may stand for the non-negative, non-positive real numbers.

$\mathbb{C}$ the set of complex numbers

$\mathbb{C}^*$ the set of non-zero complex numbers

$\overline{\mathbb{C}}$ the extended set of complex numbers $\mathbb{C} \cup \{\infty\}$ (Riemann sphere)

NAMED EQUATIONS

Abel	$f(g(x)) = f(x) + a$
Böttcher	$f(g(x)) = f(x)^p$
Cauchy I	$f(x+y) = f(x) + f(y)$
Cauchy II	$f(x+y) = f(x)\, f(y)$
Cauchy III	$f(x\, y) = f(x) + f(y)$
Cauchy IV	$f(x\, y) = f(x)\, f(y)$
conjugacy	$f(h(x)) = H(f(x))$
D'Alembert-Poisson I	$f(x+y) + f(x-y) = 2\, f(x)\, f(y)$
D'Alembert-Poisson II	$f(x+y)\, f(x-y) = f(x)^2 - f(y)^2$
Euler	$f(\lambda x, \lambda y) = \lambda\, f(x, y)$
Jensen	$f(\frac{x+y}{2}) = \frac{f(x)+f(y)}{2}$
Lobacevskii	$f(x)^2 = f(x+y)\, f(x-y)$
Pexider I	$f(x+y) = g(x) + h(y)$
Pexider II	$f(x+y) = g(x)\, h(y)$
Pexider III	$f(x\, y) = g(x) + h(y)$
Pexider IV	$f(x\, y) = g(x)\, f(y)$
Schröder	$f(g(x)) = \lambda\, f(x)$
Vincze I	$f(x+y) = g_1(x)\, g_2(y) + h(y)$
Vincze II	$f(x\, y) = g_1(x)\, g_2(y) + h(y)$
Wilson I	$f(x+y) + f(x-y) = 2\, f(x)\, g(y)$
Wilson II	$f(x+y) + g(x-y) = 2\, h(x)\, k(y)$

Bibliography

JOURNALS

[1] American Mathematical Monthly

[2] College Journal of Mathematics

[3] Crux Mathematicorum with Mathematical Mayhem

[4] Mathematics Magazine

[5] SIAM Review

[6] Journal of the Indian Mathematical Society

BOOKS

[7] J. Aczél, Lectures on Functional Analysis and Their Applications, Dover 2006.

[8] J. Aczél, J. Dhombres, Functional Equations in Several Variables, Cambridge University Press 1989.

[9] G.L. Alexanderson, L.F. Klosinski, L.C. Larson, editors, The William Lowell Putnam Mathematical Competition Problems and Solutions 1965–1984, Mathematical Association of America 1985.

[10] T. Andreescu, R. Gelca, Mathematical Olympiad Challenges, Birkahäuser 2000.

[11] T.M. Apostol, Mathematical Analysis, 2nd edition, Addison-Wesley 1974.

[12] T.M. Apostol, Introduction to Analytic Number Theory, Springer 1976.

[13] E.J. Barbeau, M.S. Klamkin, W.O. Moser, Five Hundred Mathematical Challenges, Mathematical Association of America 1995.

[14] G. Birkhoff, G.-C. Rota, Ordinary Differential Equations, 3rd edition, John Wiley & Sons 1978.

[15] A.L. Cauchy, Cours d' analyse de l' École Polytechnique, 1821.

[16] D. Djukić, V. Janković, I. Matić, N. Petrović, The IMO Compendium: A Collection of Problems Suggested for the IMOs 1959–2004, Springer 2006.

[17] H. Dörrie, 100 Great Problems of Elementary Mathematics: Their History and Solution, Dover 1965.

[18] S. Elaydi, An Introduction to Difference Equations, 3rd edition, Springer 2005.

[19] H.G. Flegg, From Geometry to Topology, Dover 1974.

[20] A.M. Gleason, R.E. Greenwood, L.M. Kelly, editors, The William Lowell Putnam Mathematical Competition Problems and Solutions 1938–1964, Mathematical Association of America 1980.

[21] S. Goldberg, Introduction to Difference Equations, Dover 1986.

[22] S.L. Greitzer, International Mathematical Olympiads 1959-1977, Mathematical Association of America 1978.

[23] G.H. Hardy, J.E. Littlewood, G. Pólya, Inequalities, Cambridge University Press 1967.

[24] H.E. Huntley. The Divine Proportion, Dover 1970.

[25] E.L. Ince, Ordinary Differential Equations, Dover 1956.

[26] K.S. Kedlaya, B. Poonen, R. Vakil, editors, The William Lowell Putnam Mathematical Competition 1985–2000, Mathematical Association of America 2002.

[27] M. Kuczma, B. Choczewski, R. Ger, Iterative Functional Equations, Cambridge University Press 1990.

[28] M. Livio, The Golden Ratio, Broadway Books 2002.

[29] B.L. Moiseiwitsch, Integral Equations, Dover 2005.

[30] D.J. Newman, A Problem Seminar, Springer-Verlag 1982.

[31] G. Pòlya, Mathematics and Plausible Reasoning, (in two volumes), Princeton University Press 1990.

[32] S. Ramanujan, Collected Papers of Srinivasa Ramanujan, edited by G.H. Hardy, P.V.S. Aiyar, and B.M. Wilson, Providence, American Mathematical Society 2000.

[33] T.L. Saaty, Modern Nonlinear Equations, Dover 1981.

[34] S. Savchev, T. Andreescu, Mathematical Miniatures, Mathematical Association of America 2003.

[35] M. Schroeder, Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise, W.H. Freeman 1991.

[36] Yu.A. Shashkin, Fixed Points, American Mathematical Society 1991.

[37] D.O. Shklarsky, N.N. Chentzov, I.M. Yaglom, The USSR Problem Book: Selected Problems and Theorems of Elementary Mathematics, Dover 1993.

[38] C.G. Small Functional Equations and How to Solve Them, Springer 2007.

[39] J. Smìtal, On Functions and Functional Equations, Taylor & Francis 1988.

[40] L. Smith, Chaos: A Very Short Introduction, Oxford University Press 2007.

[41] S.H. Strogatz, Nonlinear Dynamics and Chaos, Addison-Wesley 1994.

[42] F.G. Tricomi, Integral Equations, Dover 1985.

[43] C. Truesdell, An Essay Toward a Unified Theory of Special Functions based upon the functional equations $\partial F(z,\alpha)/\partial z = F(z,\alpha+1)$, Annals of mathematical Studies, no. 18, Princeton University Press 1948.

[44] H. Walser, The Golden Section, Mathematical Association of America 2001.

ARTICLES

[45] R.D. Boswell, *On Two Functional Equations*, Am. Math. Monthly **66** (1959) 716.

[46] D. Friedman, *The functional equation* $f(x+y) = g(x) + h(y)$, Am. Math. Monthly **69** (1962) 769.

[47] G. Hamel, *Eine basis aller zahlen und die unstetigen Lösungen der Funktionalgleichung* $f(x+y) = f(x) + f(y)$, Math. Ann. **60** (1905) 459.

[48] A. Herschfeld, *On Infinite Radicals*, Am. Math. Monthly **42** (1935) 419.

[49] E. Hille, *Functional Equations*, Chapter 1, in Lectures in Modern Mathematics, volume 3, edited by T.L. Saaty, John Wiley & Sons 1965.

[50] Pl. Kannappan, *Trigonometric Identities and Functional Equations*, The Mathematical Gazette **88** (2004) 249.

[51] T.-Y. Li, J.A. Yorke, *Period Three Implies Chaos*, Am. Math. Monthly **82** (1975) 985.

[52] R.M. May, *Simple Mathematical Models with Very Complicated Dynamics*, Nature **261** (1976) 459.

[53] P.J. McCarthy, *The General Exact Bijective Continuous Solution of Feigenbaum's Functional Equation*, Commun. Math. Phys. **91** (1983) 431.

[54] B. Reznick, *Some Thoughts on Writing for the Putnam*, in Mathematical Thinking and Problem Solving, edited by A.H. Schoenfeld, Lawrence Erlbaum Associates 1994. Reprinted in [26].

[55] S.P.S. Rathore, *On Subadditive and Superadditive Functions*, Am Math. Monthly. **72** (1965) 653.

[56] A.N. Šarkovskii, *Coexistence of Cycles of a Continuous Map of the Line into Itself*, Ukr. Mat. Ž. **16** (1964) 61. (in Russian)

[57] A.N. Šarkovskii, *On Cycles and the Structure of a Continuous Mapping*, Ukr. Mat. Ž. **17** (1965) 104. (in Russian)

[58] I.J. Schoenberg, *On the Peano Curve of Lebesgue*, Bull. Am. Math. Soc. **44** (1938) 519.

[59] W.S. Sizer, *Continued Roots*, Math. Magazine **59** (1986) 23.

[60] J. Smítal, *Chaotic Functions with Zero Topological Entropy*, Trans. AMS **297** (1986) 269.

[61] P. Štefan, *A Theorem of Šarkovskii on the Existence of Periodic Orbits of Continuous Endomorphisms of the Real Line*, Commun. Math.Phys. **54** (1977) 237.

[62] H.E. Vaughan, *Characterization of the Sine and Cosine*, Am. Math. Monthly **62** (1955) 707.

[63] M.St. Vincent, *A Functional Equation: Solution to Problem 81-1*, SIAM Review **23** (1982) 78.

Index