Preface

$$
\begin{aligned}
s^2 = &\left\{ (s-1) + \frac{1^2}{2(s-1)} {}_{+} \frac{3^2}{2(s-1)} {}_{+} \frac{5^2}{2(s-1)} {}_{+ \cdots} \right\} \\
&\times \left\{ (s+1) + \frac{1^2}{2(s+1)} {}_{+} \frac{3^2}{2(s+1)} {}_{+} \frac{5^2}{2(s+1)} {}_{+ \cdots} \right\} \qquad (1)
\end{aligned}
$$

This book has emerged as a result of my attempts to understand the theory of orthogonal polynomials. I became acquainted with this theory by studying the excellent book by Geronimus (1958). However, the fundamental reasons for its beauty and difficulty remained unclear to me. From time to time I returned to this question but the first real progress occurred only in the autumn of 1987, when I visited the Mittag-Leffler Institute right at the beginning of the foundation of the Euler International Mathematical Institute in St Petersburg (the EIMI). A simple proof of Geronimus' theorem on the parameters of orthogonal polynomials had been found (Khrushchev 1993). The paper specified an important relationship of orthogonal polynomials on the unit circle to Schur's algorithm. Later this very paper was the starting point for Khrushchev (2001). It took about eight years to complete the EIMI project, which occupied all my time, leaving no chance to continue this research.

I returned to the subject matter of this book again only in the summer of 1998 in Almaty, where I was able to get back to mathematics with the assistance of mountains and the National and Central Scientific libraries of Kazakhstan. Both these libraries had a complete collection of Euler's books translated into Russian as well as a lot of other wonderful old Russian mathematical literature such as the Russian translation of Szegő's *Orthogonal Polynomials* (1975) by Geronimus, complete with his careful and comprehensive comments. All this was very helpful for my paper Khrushchev (2001), but the question of what is the driving mechanism for, and how it explains the mystery of, orthogonal polynomials remained open. It was clear to me that most likely this mechanism is continued fractions. And my impression that something very important had disappeared from the modern theory was supported by Chebyshev's and Markoff's contributions to the subject area as well as by the following remark of Szegő in 1939: "Despite the close relationship between continued fractions and the problem of moments, and notwithstanding recent important advances in the latter subject, continued fractions have been gradually abandoned as a starting point for the theory of orthogonal polynomials".

The study of the book by W. B. Jones and W. J. Thron (1980), which I found in Pushkin's library in Almaty, indicated that perhaps the right answer to my question could be found in Euler's research on continued fractions. That not all Euler's papers on continued fractions had been carefully studied was already mentioned by Khovanskii (1957). There are two great papers of Euler in this field: (1744) and (1750b). The brief summary of Euler (1744) published as the eighteenth chapter of Euler (1748) is usually also mentioned. English translations of Euler (1744, 1748) are available. As for Euler (1750b), its first English translation is given as an appendix to this book. I thank Alexander Aptekarev (Institute of Applied Mathematics, Moscow) for arranging a translation from Latin to Russian. Then using my understanding of Euler (1750b), its Russian translation and Latin–Russian dictionaries, I translated it into English. Therefore it is not as professional as the translation of Euler (1744) but will, I hope, be acceptable. The most important facts on continued fractions from Euler (1744, 1748) are presented in the first chapter here.

I saw the Latin version of Euler (1750b) only in January 2003 when, at the kind invitation of Barry Simon, I visited Caltech, California, to lecture from a preliminary version of the present book. At that time I knew nothing of the project of the Euler Archive run by the Euler Society (www.eulersociety.org), which possibly then was only under construction. Lecturing in front of Simon's group in Caltech strongly reminded me of the golden years in St Petersburg in the 1970s. There are of course some differences because of the location. For instance, they do not have late evening tea and instead take lunch before the seminar where you can enjoy, if you are brave enough, hot Mexican pepper.

Even a very brief inspection of Euler (1750b) shows that it was motivated by the remarkable formula (1) discovered in March 1655 by Brouncker, the first President of the Royal Society of London. The proof of this formula was included in section 191 of Wallis (1656). I did not have access to this striking book at that period but from historical literature in Russian, for instance from Kramer (1961), I discovered that the presentation in this particular part is just impossible to understand. This was indirectly confirmed by Euler (1750b), who, in spite of the fact that Wallis' *Arithmetica Infinitorum* was a permanent feature of his desk, complained that Brouncker's proof was seemingly irreparably lost.

Nonetheless, in the summer of 2004 in the mountains of Almaty I came to the conclusion that possibly this can be done very easily if suitably transformed partial Wallis products are written as continued fractions. I couldn't find the required transformation and was about to give up but suddenly help arrived from the Amazon bookstore. It has a very good knowledge system which makes proposals based on the captured interests of their customers. Amazon's email claimed that the English translation of *Arithmetica Infinitorum* by Jacqueline Stedall (2004) was available from Oxford.

When I arrived back in Ankara the book awaited me in the post office. I opened Wallis' comment on section 191 and saw the following: "*The Noble Gentleman noticed that two consecutive odd numbers, if multiplied together, form a product which is the square of the intermediate even number minus one... He asked, therefore, by what ratio the factors must be increased to form a product, not those squares minus one, but equal to the squares themselves*". When I read this I could immediately understand how Brouncker proved (1). It took some time to complete the calculations and this proof is now available in Chapter 3. Wallis' previously unclear remarks are now used to confirm that the proof presented is exactly that discovered by Brouncker.

A few words explaining why (1) is so important. It is the functional equation $b(s-1)b(s+1) = s^2$, reminding us on the one hand of an elementary formula $(s-1)(s+1) = s^2 - 1$ from algebra and, on the other hand, of the functional equation for Euler's gamma function $\Gamma(x+1) = x\Gamma(x)$. In fact these two functions are related by the Ramanujan formula (see Theorem 3.25). Another mystery is that Ramanujan's formula in turn is an easy consequence of Brouncker's theory... Combining the Ramanujan formula with Chebyshev's arguments presented in Section 7.4, one easily obtains that the polynomials written explicitly in Wallis (1656, §191) are orthogonal with respect to the weight

$$d\mu = \frac{1}{8\pi^3}\left|\Gamma\left(\frac{1+it}{4}\right)\right|^4 dt\,.$$

In (1977), J. A. Wilson, following some ideas of R. Askey on the gamma function, introduced a new class of orthogonal polynomials depending on a number of independent parameters. An impressive property of Wilson's polynomial family is that almost all the so-called classical orthogonal polynomials are placed on its boundary. An inquiry into Andrews, Askey and Roy (1999) shows that, on the contrary, Brouncker's polynomials are placed at the very center, corresponding to the choice $a = 0$, $b = 1/2$, $c = d = 1/4$. Thus Brouncker's formula in 1655 already listed important orthogonal polynomials, though not in a direct form. But neither was the Universe in its first few minutes similar to the present world. In addition to special functions, Brouncker's formula stimulated, or it is better to say could stimulate, developments in two other important directions.

The first is the moment problem considered by Stieltjes (1895). One can easily notice a remarkable similarity of Brouncker's arguments to those of Stieltjes. The second is the solution of Pell's equation obtained by Brouncker as his answer to the challenge of Fermat. It looks as if Fermat carefully studied Wallis' book. Still, I have never heard that he ever mentioned §191 in his letters. Instead Fermat proposed to outstanding British mathematicians a problem which they could solve by the method of Brouncker presented in this very paragraph. And indeed Brouncker solved Fermat's problem by applying a part of the argument he used to answer Wallis' question. After that Wallis developed his own method. This is considered in more detail in Chapter 2.

It is just unbelievable that such a partial, on first glance, result obtained in 1655 encapsulated a considerable part of the further development of algebra and analysis. True, this was a result on the quadrature problem obtained with continued fractions...

From the critical analysis of Brouncker's proof of (1), two interesting properties of continued fractions can be observed. If some continued fractions give a development of one part of mathematics then it is quite possible that similar progress can be made with other continued fractions in another part. In most cases the arguments could be simplified, as for instance Wallis did for Pell's equation, but at the cost of losing some substantial relationships, regarding which Euler was such a great master. I assume that the right explanation of this phenomenon lies in approximation theory. Any continued fraction is nothing other than an algorithm whose elementary steps are simple Möbius transforms. Therefore, adjusting these parameters in an appropriate way at each step, one can significantly change the original result. The art is to make this choice properly so that a new result can at least be stated.

In 1880 A. A. Markoff completed his Master's thesis at St Petersburg University, which was devoted to the theory of binary quadratic forms of positive determinant. I strongly believe that this was the best work of Markoff's whole mathematical career. It appears to have determined his later significant papers, in particular those in probability theory. It was not just one more application of continued fractions. Rather it was an incredibly beautiful demonstration of what can be done with their proper use. Therefore, although I was forced to sacrifice Stieltjes' theory of moments to a great extent in consequence, I have included this important theory in Chapter 2. The theory of moments is well presented in a number of books (Akhiezer 1961, Shohat and Tamarkin 1943 and Stieltjes 1895), whereas Markoff's original approach to this problem is not. In addition Markoff's theory has some relations to my own research (Khrushchev 2001a, b, 2002).

The key to both is Lagrange's formula (1.50). The Lagrange function $\mu(\xi)$ is defined for irrational ξ as the supremum of $c > 0$ such that

$$\left|\frac{p}{q} - \xi\right| < \frac{1}{cq^2}$$

has infinitely many solutions in the integers p, q, $q > 0$. So, the greater $\mu(\xi)$ is, the better can ξ be approximated by rational numbers. The range of μ is called the Lagrange spectrum. Markoff proved that, on the one hand, for $\mu(\xi) < 3$ the Lagrange spectrum is discrete and any ξ with $\mu(\xi) < 3$ is a quadratic irrational that is equivalent to the continued fraction

$$\xi(\theta, \delta) = \frac{1}{r_1} + \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_2} + \frac{1}{r_3} + \frac{1}{r_3} + \cdots + \frac{1}{r_n} + \frac{1}{r_n} + \cdots, \tag{2}$$

with $\delta = 0$ and a rational $0 < \theta \leqslant 1$. Here

$$r_n = [(n+1)\theta + \delta] - [n\theta + \delta]$$

is a Jean Bernoulli sequence,[1] which Jean Bernoulli introduced in his treatise on astronomy (1772). On the other hand, $\mu(\xi(\theta, \delta)) = 3$ for irrational $\theta \in (0, 1)$. Hence there are transcendental numbers with $\mu(\xi) = 3$. Moreover, they may be represented by regular continued fractions (2), which are simply expressed via Jean Bernoulli sequences. These are the worst transcendental numbers from the point of view of rational approximation, as follows from Markoff's main result, in contrast with Liouville's constant

$$\begin{aligned} L &= \sum_{n=0}^{\infty} 10^{-n!} \\ &= \frac{1}{9} + \frac{1}{11} + \frac{1}{99} + \frac{1}{1} + \frac{1}{10} + \frac{1}{9} + \frac{1}{999\,999\,999\,999} + \frac{1}{1} + \cdots. \end{aligned}$$

In Khrushchev (2001, 2002), Lagrange's formula, see Theorem 8.67, is applied not to regular continued fractions but to Wall continued fractions, which are nothing other than a form of the classical Schur algorithm. In the case of numbers one usually considers either their decimal representations or regular continued fraction expansions, and in this case there are three closely related objects. The first is the continuum $\mathfrak{P}(\mathbb{T})$ of all probability Borel measures on $\mathbb{T}$. The second and the third are the continuums of analytic functions F^σ with positive real part in the unit disc $\mathbb{D}$:

$$F^\sigma(z) = \int_{\mathbb{T}} \frac{\zeta + z}{\zeta - z} d\sigma(\zeta) = \frac{1 + z f^\sigma}{1 - z f^\sigma} \tag{3}$$

and contractive analytic functions f^σ in $\mathbb{D}$. Any such f^σ expands into a Wall continued fraction

$$f(z) = a_0 + \frac{(1 - |a_0|^2)z}{\bar{a}_0 z} + \frac{1}{a_1} + \frac{(1 - |a_1|^2)z}{\bar{a}_1 z} + \cdots, \tag{4}$$

where, by Geronimus' theorem, which I mentioned right at the start, $\{a_n\}_{n\geqslant 0}$ are on the one hand the Verblunsky parameters of σ and on the other hand the Schur parameters of f^σ. The even convergents to (4) are contractive rational functions A_n/B_n, which by Schur's theorem (1917) converge to f^σ uniformly on compact subsets of $\mathbb{D}$. The substitution of f^σ in (3) with A_n/B_n results in rational functions Ψ_n^*/Φ_n^*, where $\{\Phi_n\}_{n\geqslant 0}$ are monic orthogonal polynomials in $L^2(d\sigma)$. By Lagrange's formula, asymptotic

[1] So-called by A. A. Markoff, who discovered important properties of these sequences.

properties of the normalized orthogonal polynomials $\{\varphi_n\}_{n\geqslant 0}$, $|\varphi|^2 d\sigma \in \mathfrak{P}(\mathbb{T})$, can be studied from the point of view of approximation on $\mathbb{T}$ either of f^σ by A_n/B_n or F^σ by Ψ_n^*/Φ_n^*. For instance, it turns out that Szegő measures, i.e. measures with finite entropy

$$\int_{\mathbb{T}} \log \sigma' dm > -\infty \,, \tag{5}$$

where m is the Lebesgue measure on $\mathbb{T}$, are exactly the measures such that $A_n/B_n \to f$ in L^1, where the distance between values of A_n/B_n and f is measured in the Poincaré metric of the non-euclidean geometry of $\mathbb{D}$; see Theorem 8.56.

A measure $\sigma \in \mathfrak{P}(\mathbb{T})$ is called a *Rakhmanov* measure if

$$*\text{-}\lim_n |\varphi_n|^2 d\sigma = dm$$

in the $*$-weak topology of $\mathfrak{P}(\mathbb{T})$. A measure is a Rakhmanov measure if and only if the Máté–Nevai condition

$$\lim_n a_n a_{n+k} = 0 \ \text{ for } \ k \geqslant 1 \,,$$

for the Verblunsky parameters $\{a_n\}_{n\geqslant 0}$ is satisfied (Theorem 8.73). Moreover $A_n/B_n \Rightarrow f^\sigma$ in measure on $\mathbb{T}$ for any Rakhmanov measure σ. With this theorem to hand we can prove that $A_n/B_n \Rightarrow f^\sigma$ in measure on $\mathbb{T}$ if and only if either σ is singular or $\lim_n a_n = 0$; see Theorem 8.78. The last two results have an important practical application. Let σ be any Szegő measure. By Geronimus, theorem (5) is equivalent to

$$\sum_{n=0}^{\infty} |a_n|^2 < \infty \,.$$

Let us now modify the sequence $\{a_n\}_{n\geqslant 0}$ on an arbitrary sparse subset Λ of integers, which in addition starts far from $n = 0$. We replace a_n with $10^{-1000!}$ if $n \in \Lambda$. The sequence obtained, $\{a_n^*\}_{n\geqslant 0}$, is a Máté–Nevai sequence. Hence the measure σ^* with Verblunsky parameters $\{a_n^*\}_{n\geqslant 0}$ is a Rakhmanov measure, implying that $A_n/B_n \Rightarrow f^{\sigma^*}$ on $\mathbb{T}$. Since $\lim_n a_n^* \neq 0$, we obtain that σ^* is singular. It is impossible to distinguish in practice the Szegő measure σ from the singular measure σ^* just by observing their first Verblunsky parameters.

We also construct in Chapter 8 examples of extremely transcendental σ, such that the sequence $\{|\varphi_n|^2 d\sigma\}_{n\geqslant 0}$ is dense in $\mathfrak{P}(\mathbb{T})$. Elements of the theory of periodic measures are also considered at this point.

Although the analogy described above is rather remote, in my opinion it is beautiful and justifies the inclusion of Markoff's results in Chapter 2. At many points I follow the Russian version of Markoff's thesis. However, there is an important difference. From Lagrange's formula Markoff quite naturally arrives at a combinatorial description of Jean Bernoulli sequences. In this book the properties of Jean Bernoulli sequences are first studied in detail in Chapter 1. Then these results are applied in Chapter 2, which makes the basic ideas of Markoff look more natural. With Jean Bernoulli sequences and the formulas for Jean Bernoulli and Markoff periods one can easily calculate numerically as many points of the Lagrange spectrum as necessary.

Markoff's theory completes the algebraic part of this book. The analytic part begins in Chapter 3 and is followed in Chapter 4 by Euler's research. To a great extent this chapter covers Euler (1750b) but with the difference that Brouncker's method recovered in Chapter 3 is applied. The method can be extended from the unit circle considered by Wallis to the class of sinusoidal spirals introduced into mathematics in 1718 by another great British mathematician, Colin Maclaurin. Chapter 4 covers the forgotten Euler differential method of summation of some continued fractions of hypergeometric type. It turns out that, for instance, approximately half the continued fractions discovered later by Stieltjes can be easily summed up by this method of Euler. However, the differential method does have some limitations. Attempting to overcome them Euler arrived at the beautiful theory of Riccati equations. In (1933) Sanielevici presented Euler's method in a very general form. Later Khovanskii (1958) using Sanielevici's results developed as continued fractions many elementary functions. Still, I think that Euler's method as stated by Euler makes everything more clear. In this part I filled some gaps in the proofs while trying not to violate Euler's arguments. The central result here is the continued fraction for the hyperbolic cotangent. I collect in Chapter 5 some results which were or could be directly or indirectly influenced by Euler's formulas. Such an approach sheds new light on the subject.

Chapter 6 presents results either obtained by Wallis interpolation or by a direct transfer from the regular continued fractions of number theory to polynomial continued fractions, i.e. to P-fractions. Euler's results on hypergeometric functions play a significant role here. Another interesting topic of this chapter is the periodicity of P-fractions. As before, the first results were obtained by Euler. Using continued fractions of the radicals of quadratic polynomials as guidance, Euler found his now well-known substitutions for integration. This was extended by Abel in one of his first papers, which incidentally preceded his discoveries in elliptic functions. I include in this chapter a beautiful result of Chebyshev on integration in finite terms.

Chapter 7 indicates how Euler's ideas eventually led to the discovery of orthogonal polynomials. Finally, I present in Chapter 8 my own research on the convergence of Schur's algorithm.

A few words on the title. It varied several times but since the essential part of the book is related to Euler I believe that finally I made a good choice. Moreover, in 2007 the tercentenary of Euler was celebrated.

Following Euler, I have split the book into small numbered "paragraphs" (subsections). This was an old tradition in mathematics, now almost forgotten. It makes the book easier to read. The difference from what Euler did is that almost all the "paragraphs" also have titles.

This book is not a complete account of what has been done in orthogonal polynomials or in continued fractions. In orthogonal polynomials Szegő's book (1939) is still important. There are also two important contributions made by Nevai (1979, 1986), and another two books by Saff and Totik (1997) and by Stahl and Totik (1992). As to orthogonal polynomials on the unit circle there is the recent and exhaustive work in two volumes of Barry Simon (2005). More on continued fractions can be found in Jones and Thron (1980), Khinchin (1935), Khovanskii (1958), Perron (1954, 1957) and Wall (1943).

Most parts of this book require only some knowledge of calculus and an undergraduate course on algebra. In Chapters 6–8 elementary facts from complex analysis are used occasionally. In Chapter 8 in addition it is important to know basic facts on Hardy spaces. There are two relatively new and very well written books on Hardy spaces: Garnett (1982) and Koosis (1998).

I wish to thank a number of people and organizations supporting me in one or another way during the work on this project. First of all I express deep gratitude to my aunt Galina Kreidtner. She was not a mathematician, she was an architect and artist. Nonetheless she enthusiastically and helpfully discussed with me the idea of this book in Almaty. She died at the age of 80 in 2000 and so I lost one of the best friends in my life. I dedicate this book to her memory.

Another very good friend, Paul Nevai from the Ohio State University, made a right choice in favor of orthogonal polynomials at the very beginning of his mathematical career in St Petersburg University, where we were fellow students. His support was also extremely valuable and sincere.

My special thanks to Purdue University, Indiana, which played a significant role in my career in mathematics at least twice. I am particularly grateful to David Drasin and Carl Cowen.

I am very grateful also to Atilim University, Ankara, which has created very good conditions for my research in mathematics and supported all my scientific projects.

I express my sincere and deep gratitude to Barry Simon (Caltech, Pasadena) who provided very important personal support for this book.

The book turned out to be so much influenced by Wallis' *Arithmetica Infinitorum* (1656) that I cannot avoid a temptation to finish this preface with a citation from its very end:

> *There remains this: we beseech the skilled in these things, that what we thought worth showing, they will think worth openly receiving, and whatever it hides, worth imparting more properly by themselves to the wider mathematical community.*

PRAISE BE TO GOD

Almaty–Ankara (1998–2007)