

In Praise (and Search) of J. V. Uspensky

Persi Diaconis and Sandy Zabell

Abstract. The two of us have shared a fascination with James Victor Uspensky’s 1937 textbook *Introduction to Mathematical Probability* ever since our graduate student days: it contains many interesting results not found in other books on the same subject in the English language, together with many non-trivial examples, all clearly stated with careful proofs. We present some of Uspensky’s gems to a modern audience hoping to tempt others to read Uspensky for themselves, as well as report on a few of the other mathematical topics he also wrote about (e.g., his book on number theory contains early results about perfect shuffles).

Uspensky led an interesting life: a member of the Russian Academy of Sciences, he spoke at the 1924 International Congress of Mathematicians in Toronto before leaving Russia in 1929 and coming to the US and Stanford. Comparatively little has been written about him in English; the second half of this paper attempts to remedy this.

Key words and phrases: J. V. Uspensky, A. A. Markov, probability theory, Markov’s method of continued fractions, Bernoulli’s theorem, Lexis ratio, card shuffling, Russian mathematics, Stanford Mathematics Department.

1. INTRODUCTION

In 1927, when Harald Cramér visited England and mentioned to G. H. Hardy (his friend and former teacher) he had become interested in probability theory, Hardy replied “there was no mathematically satisfactory book in English on this subject, and encouraged me to write one” (Cramér, 1976, p. 561). Ten years later, J. V. Uspensky’s book *Introduction to Mathematical Probability* (1937) filled this vacuum: “his mathematically demanding book was the standard text on probability theory in [the US] until the appearance” of Feller’s 1950 classic, *Introduction to Probability Theory and its Applications* (Reeds, Diffie and Field, 2015, p. 588).¹

Persi Diaconis is the Mary V. Sunseri Professor of Statistics and Mathematics, Stanford University, Stanford, California 94305, USA (e-mail: diaconis@math.stanford.edu). Sandy Zabell is Professor, Departments of Mathematics and Statistics, Northwestern University, Evanston, Illinois 60208, USA (e-mail: zabell@math.northwestern.edu).

¹See also Seneta (2006, p. 6) (“Uspensky’s book seems to have brought analytical probability, in the St. Petersburg tradition, to the United States, where it remained a primary probabilistic source until the appearance of W. Feller’s *An Introduction to Probability Theory and Its Applications* in 1951”); Bhatia (2008, p. 26), quoting S. R. S. Varadhan (“There were not too many books available at that time [the mid-1950s]. Feller’s book had just come out. Before that there was a book by Uspensky. These were the only books on Probability”). For a contemporary reference, see Munroe (1951, p. 2).

Who was the author of this important book? J. V. (James Victor) Uspensky (born Yakov Viktorovich Uspensky, Яков Викторович Успенский, 1883–1947) entered the Saint Petersburg Imperial University in 1903, received his undergraduate degree there in 1906 as well as his graduate degree in 1911, studying under the great Andrei Andreievich Markov (1856–1922). In 1915, he became a Professor at Petrograd (the new name for Saint Petersburg) University, and in 1921 elected a member of the Russian Academy of Sciences. He gave a talk at the 1924 International Congress of Mathematicians, and traveled to the United States several times during this period: 1924, 1926–1927 (when he taught at Carleton College in Minnesota), and 1929, when he moved permanently to the US and Stanford. Initially appointed an acting Professor of Mathematics (1929–1931), he became a permanent member of Stanford’s mathematics department in 1931 and remained there until his death in 1947. He is best known today for three textbooks he wrote during his time at Stanford although, as we shall see, he was much more active than just this.

In the following, we focus on Uspensky’s three textbooks (1937, 1939, 1948), primarily the one on mathematical probability, but then circle back and take a closer look at this very interesting person. While this paper is intended primarily as a contribution to the history of our subject, it is worth mentioning that we have used some of the topics discussed below as course material in both introductory and graduate courses. Historical material often

brings a course to life for students. The current proofs in textbooks are frequently so streamlined it can be easy to forget just how challenging the initial proofs were. Some of the material in the sections below can be used directly as part of a lecture, or the basis of a challenging problem set, or as reading in conjunction with a student project.

2. INTRODUCTION TO MATHEMATICAL PROBABILITY

In the preface to his book, Uspensky briefly explained its origin. It grew out of a course on mathematical probability Uspensky had taught at Stanford for a number of years. At first he prepared a short mimeographed book covering only the elementary aspects of the subject, but over time the scope of the course enlarged and in the end “a rather large manuscript” covering the most important parts of the theory emerged. Originally, Uspensky did not intend to publish this, but did so after the urging of students and other readers.

Uspensky himself saw the book as dividing naturally into two parts: the first, “consisting of Chapters I to XII inclusive, is accessible to a person without advanced mathematical knowledge”; Chapters XIII through XVI, in contrast, “incorporate the results of modern investigations,” and as a result were necessarily “more complex” and required “more mature mathematical preparation” (covering distribution theory, rigorous proof of the central limit theorem, and mathematical statistics).

There were, of course, other expositions of the mathematical theory then available in English, notably Fisher (1922) and Fry (1965). Both of these however were very different from Uspensky’s in both their aim and scope. Arne Fisher (1887–1944) was an applied statistician who worked first at Prudential Insurance and then Western Union Telegraph. Although interested in the philosophical underpinnings of mathematical probability, the great strength of his book lies in its discussion of the Lexis ratio and the work of the Scandinavian school of statistics.² Thornton Carl Fry (1882–1991) likewise came from an applied background, working at Western Electric and Bell Telephone. His book is very interesting, but much of it is devoted to applied probability and it does not approach Uspensky’s in either mathematical depth or breadth.³

Uspensky’s book was well received.⁴ One of its great strengths was its making available in English for the first

time a substantial fraction of the Russian literature on probability. In the following, we discuss a few particularly interesting examples.

2.1 Computing Binomial Tail Probabilities

Suppose S_n is the number of heads in n independent tosses of a p -coin. If $n = 9000$ and $p = 1/3$, the normal approximation tells us that $P(S_{9000} > 3090)$ is about 0.02209. The R statistical software package tells us the exact value to five places is 0.02170, so the normal approximation here is accurate to three places. But how can we tell this without R? Estimating the error in the normal approximation or doing better by adding correction terms to it can be tricky. It turns out there is a very clever way of tackling the twin issues of better approximation and error estimation without appealing at all to the central limit theorem (or brute force calculation).

In Chapter 3, Section 8 of his book, Uspensky notes that although there are asymptotic formulas for estimating interval binomial probabilities of the form $P(a < S_n < b)$, “less known is the ingenious use by Markoff of continued fractions for that purpose.” In the following we give (1) Markov’s method as described by Uspensky, together with a proof of the key inequalities; (2) a discussion of the method, including an important forward recursion the convergents satisfy permitting their efficient computation and (3) brief comments on some later literature.

2.1.1 Markov’s method of continued fractions. Let $b(k; n, p)$ be the binomial probability of k successes in n trials with success probability p , let l be an integer greater than np , and consider the right binomial tail probability

$$P(l) := P(S_n > l).$$

Then a trite calculation yields

$$P(l) = b(l+1; n, p)S$$

where

$$S = 1 + \frac{n-l-1}{l+2} \frac{p}{q} + \frac{(n-l-1)(n-l-2)}{(l+2)(l+3)} \left(\frac{p}{q}\right)^2 + \dots$$

The sum S is a special case of the hypergeometric series

$$F(\alpha, \beta, \gamma, x) = 1 + \frac{\alpha\beta}{1 \cdot \gamma} x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{1 \cdot 2\gamma(\gamma+1)} x^2 + \dots;$$

namely

$$F\left(-n+l+1, 1, l+2, -\frac{p}{q}\right).$$

² “[Arne] Fisher’s greatest contribution was calling to the attention of American mathematicians the work of the great Scandinavian statisticians. Gram, Thiele, Westergaard, Charlier, were mere names to us before the publication of Fisher’s *The Mathematical Theory of Probabilities* in 1922” (Molina, 1944, p. 251)

³ Two other textbooks then available were Coolidge (1925) and Burnside (1928). Both were much inferior to Uspensky’s.

⁴ David (1938), for example, in her review of it in *Biometrika* wrote “This book is so good that it should remain a classic in the literature of the theory of probability for many years”.

A hypergeometric series of this type has a continued fraction expansion that is a special case of Gauss's continued fraction for a ratio of two hypergeometric functions (for the latter, see Wall, 1948). In this particular case, if

$$c_k = \frac{(n-k-l)(l+k)}{(l+2k-1)(l+2k)} \frac{p}{q},$$

$$d_k = \frac{k(n+k)}{(l+2k)(l+2k+1)} \frac{p}{q},$$

then the resulting expansion is

$$S = \frac{1}{1 - \frac{c_1}{1 + \frac{d_1}{1 - \frac{c_2}{1 + \ddots - \frac{c_{n-l-1}}{1 + \frac{d_{n-l-1}}{1}}}}}}}$$

For each $k \geq 1$, let

$$C_k = \frac{1}{1 - \frac{c_1}{1 + \frac{d_1}{1 - \ddots - c_k}}}$$

and

$$D_k = \frac{1}{1 - \frac{c_1}{1 + \frac{d_1}{1 - \ddots - \frac{c_k}{1 + d_k}}}}}$$

We will refer to the C_k and D_k as “ C -convergents” and “ D -convergents.” It is not hard to see that the Markov method brackets S by a C -convergent and a D -convergent. In particular, it is not hard to show the convergents exhibit a two-step ping-pong type of behavior, satisfying the successive sequence of inequalities

$$C_2 < D_2 < C_4 < D_4 < \cdots < S$$

$$< \cdots < D_3 < C_3 < D_1 < C_1$$

(until $k = n - l - 1$, when $D_k = S$).

Because it is surprisingly hard to find an actual proof of the alternating behavior of the convergents in the literature, we give the details here.

PROOF. Because $l > np$, is easy to see that

$$1 > c_1 > \cdots > c_k > \cdots > c_{n-l} = 0.$$

Let

$$\omega_k := \frac{c_k}{1 + \frac{d_k}{1 - \frac{c_{k+1}}{1 + \ddots}}}$$

Note that by definition one has that

$$S = \frac{1}{1 - \omega_1},$$

as well as the recursive relation

$$\omega_k = \frac{c_k}{1 + \frac{d_k}{1 - \omega_{k+1}}}.$$

Furthermore, it is easily seen that $0 < \omega_k < c_k$ for $k < n - l$.

The proof proceeds by exploiting the recursive nature of the ω_k . Note that for $0 < \alpha < \beta$,

$$\frac{c_k}{1 + \frac{d_k}{1 - \beta}} < \frac{c_k}{1 + \frac{d_k}{1 - \alpha}}.$$

Call this the *basic inequality*. Because $0 < \omega_{k+1} < c_{k+1}$, it follows from the basic inequality and the ω_k recursion that

$$\frac{c_k}{1 + \frac{d_k}{1 - c_{k+1}}} < \omega_k < \frac{c_k}{1 + d_k} < c_k.$$

If one then continues this process, thus extending the continued fraction back by the additional two terms c_{k-1} and d_{k-1} , the result is a new set of inequalities with the direction reversed. (Note this is an immediate consequence of now applying the basic inequality three times.) Continuing this a total of k times, going all the way to c_1, d_1 and then finally passing from bounds for ω_1 to bounds for $S = 1/(1 - \omega_1)$ (note this last step does not reverse the inequalities), one ends up with

$$(1) \quad (-1)^k C_{k+1} < S < (-1)^k D_k < (-1)^k C_k.$$

Furthermore, since

$$\frac{c_{k+1}}{1 + \frac{d_{k+1}}{1 - c_{k+2}}} > 0,$$

once again invoking the basic inequality and arguing as before, one has

$$(2) \quad (-1)^{k+2} C_{k+2} < (-1)^k D_k.$$

It follows from (1) and (2) that the convergents satisfy the indicated sequence of inequalities. \square

Uspensky illustrates this computational process in detail for the case when $n = 9000$, $p = 1/3$ and $l = 3090$, and starting at ω_6 obtains the bounds

$$0.02161 < P(S_{9000} > 3090) < 0.02175.$$

As noted earlier, the exact answer to five places is 0.02170, so in this example the continued fraction approximation is accurate to three places.

2.1.2 Discussion. Although Markov's method is at once both elegant and useful, there are some puzzling aspects to Uspensky's presentation. First, Uspensky computes the convergents of the continued fraction "from the bottom up." This is fine as far as it goes, but inefficient if one then decides to compute further out in the continued fraction because you have to start over from scratch.

Following the notation in [Dudley \(1987\)](#), consider the interweaved set of convergents $Q_1 = 1$ and

$$Q_{2k} = C_k, \quad Q_{2k+1} = D_k, \quad k \geq 1.$$

Let $Q_n = A_n/B_n$ (A_n and B_n relatively prime). The numerators A_n and denominators B_n of the convergents of an alternating continued fraction satisfy a simple two-step linear recursion, which facilitates their computation; see [Dudley \(1987, p. 589, Theorem 3.1\)](#). Here, the formula for the recursion takes the form: if X_n represents either the numerator A_n or the denominator B_n , then for $k \geq 1$,

$$X_{2k} = X_{2k-1} - c_k X_{2k-2},$$

$$X_{2k+1} = X_{2k} + d_k X_{2k-1},$$

subject to the initial conditions $A_0 = 0, A_1 = 1, B_0 = B_1 = 1$. One can thus readily compute successive convergents using the recursive formula, knowing S is always bounded by two successive convergents D_k, C_{k+1} lying on either side of S , stopping when a desired degree of accuracy is reached. It is unclear why Uspensky does not mention this, since he was certainly familiar with this phenomenon; see page 359 of his own book.

For some reason, Uspensky does not give a specific reference for Markov's method. It in fact appears in Markov's 1900 book *Ischislenie veroiatnostei* (Calculus of Probabilities), Section 24, pages 150–157; and Uspensky's treatment follows that of Markov's very closely, including the layout of the illustrative numerical example (although Uspensky uses different values for n, p and l). Markov does note (p. 153) the odd C -convergents are greater than S and the even ones less, but does not mention their monotonic nature and says nothing about the D -convergents. It is hard to believe however he did not know the full story. Perhaps, focusing on the computation for a single value of k , he did not think it pedagogically desirable to go into such details. For discussion of Markov's treatment, see [Dutka \(1981, pp. 19–21\)](#), [Seneta \(2013, pp. 1105–1106\)](#).

The use of continued fractions in probability to perform computations of the normal integral goes back to Laplace ([Hald, 1998, pp. 208–209](#)). Chebyshev later used them to study least squares interpolation ([Hald, 1998, pp. 525–527](#)), as well as part of a theoretical attack on the moment problem. Markov's 1884 thesis ("On some applications of algebraic continued fractions"), continued this work of his advisor Chebyshev; see [Uspensky \(1937, Appendix II\)](#). Markov also published a short book in French shortly after ([Markov, 1888](#)) in which he used continued fractions to compute tables of the normal integral.

2.1.3 Later literature. Markov's book went through four editions (1900, 1908, 1913, 1924), but the section on the method of continued fractions in the 1924 edition (Section 18, pp. 104–114), is virtually identical to that in the 1900 edition. The second edition is conveniently available in German translation ([Liebmann, 1912](#)); see Section 25, pages 135–141 for the material on continued fractions.

Uspensky appears initially to have been virtually the sole source—perhaps the sole source—for discussion of the method in the English literature. [Bahadur \(1960\)](#) independently came up with a closely related representation of the tail as a product of the lead term and q times a different hypergeometric series,

$$P(S_n \geq j) = \left[\binom{n}{j} p^j q^{n-j} \right] q F(n+1, 1; j+1; p).$$

(The simple argument invokes a standard relation between the binomial and negative binomial distributions.) Bahadur then goes on to add:

Another method of using continued fractions to obtain bounds on B , which is based [directly on the tail sum] itself rather than [a series derived using the negative binomial distribution], is given in Uspensky ([2], pp. 52–56). This method, which is attributed in [2] to Markov, does not appear to be generally known, and might therefore be described here.

Much of the subsequent literature on Markov's method is curiously terse when it comes to the monotonic nature of the convergents. Bahadur says "it can then be shown," citing Uspensky, but gives no page reference and, as discussed above, Uspensky is in fact silent on this issue. [Dudley \(1987, p. 589\)](#) passes over both Bahadur and Uspensky in silence and directly cites Markov (1924, p. 108). He appears to have felt some frustration with the discussion in the literature about the two-step alternating nature of the convergents, for he says of it that it is "known, at least in some cases," citing both Markov (1924, p. 108) and [Peizer and Pratt \(1968, p. 1452\)](#). He says it "follows directly from" the generalized continued fraction representation.

For a more general discussion of continued fraction expansions for hypergeometric series, see [Borwein, Choi and Pigulla \(2005\)](#).

2.2 Bernoulli's Theorem

Nowadays the weak law of large numbers for independent and identically distributed trials is usually proven using Chebyshev's proof. But this is not the one Uspensky gives for Bernoulli trials. Instead he says:

Several proofs of this important theorem are known which are shorter and simpler but less

natural than Bernoulli's original proof. It is his remarkable proof that we shall reproduce here in modernized form.

2.2.1 Bernoulli's proof. The strategy underlying James Bernoulli's original proof is straightforward: divide the range of the binomial tail probability into blocks, bound the sum of the terms in each block by a corresponding term in a geometric series, and then show the sum of the resulting geometric series satisfies the requisite inequality for all n sufficiently large. The actual tactical execution of this however requires considerable skill. Here is an outline of the proof Uspensky gives using his notation.

We seek to show that if there are m heads in n tosses of a p -coin, then for any $\epsilon, \eta > 0$, one has

$$P\left(\left|\frac{m}{n} - p\right| < \epsilon\right) > 1 - \eta$$

for all n sufficiently large.

Here is the proof. Let $\lambda = \lceil np \rceil$, $\mu = \lceil np + n\epsilon \rceil$ denote respectively the ceiling of np and $np + n\epsilon$, respectively; that is, the integers λ, μ such that

$$\lambda - 1 < np \leq \lambda, \quad \mu - 1 < np + n\epsilon \leq \mu.$$

Let $g := \mu - \lambda$ (this will be the block size), let T_k be the probability of k heads in n tosses of a p -coin, and set

$$\begin{aligned} A &= T_\lambda + \cdots + T_{\mu-1}, \\ C &= T_\mu + \cdots + T_n, \\ A_j &= T_{\mu+(j-1)g} + T_{\mu+(j-1)g+1} \\ &\quad + \cdots + T_{\mu+jg-1}, \quad j \geq 1. \end{aligned}$$

so that

$$C = A_1 + A_2 + A_3 + \cdots.$$

The proof then breaks down into the following steps:

1. *Bound the tail by a geometric series.* For integers $b > a \geq 0$ and $k > 0$, a simple manipulation of inequalities gives

$$\frac{T_{b+k}}{T_b} < \frac{T_{a+k}}{T_a};$$

and it is easily seen from this (setting $A_0 := A$) that

$$A_j < A_{j-1} \left(\frac{T_\mu}{T_\lambda} \right), \quad j \geq 1$$

and, therefore,

$$C < A \left[\left(\frac{T_\mu}{T_\lambda} \right) + \left(\frac{T_\mu}{T_\lambda} \right)^2 + \left(\frac{T_\mu}{T_\lambda} \right)^3 + \cdots \right].$$

2. *Derive a bound for the common ratio of this series.* For $x \geq \lambda \geq np$, it is easily seen that $T_{x+1}/T_x < 1$. Expressing T_μ/T_λ as a telescoping product of g terms, retaining only the first $\alpha \leq g$ of these terms (i.e., the length

of some subblock), invoking the first inequality in Part 1 above, and noting that

$$\frac{T_{\mu-\alpha+1}}{T_{\mu-\alpha}} \leq \left(\frac{n - \mu + \alpha}{n - \alpha + 1} \frac{p}{q} \right),$$

gives

$$\frac{T_\mu}{T_\lambda} < \left(\frac{n - \mu + \alpha}{n - \alpha + 1} \frac{p}{q} \right)^\alpha.$$

3. *Use the preceding to show the common ratio of the series is bounded by η .* So far $\alpha \leq g$ can be any sub-block size. Now choose α to be the least positive integer such that

$$\left(\frac{p}{p + \epsilon} \right)^\alpha \leq \eta.$$

Further manipulation of inequalities shows that for

$$n \geq \frac{\alpha(1 + \epsilon) - q}{\epsilon(p + \epsilon)}$$

one has both $g \geq \alpha$ (so that the last inequality in the previous step holds) and

$$\frac{n - \mu + \alpha}{n - \alpha + 1} \frac{p}{q} < \frac{p}{p + \epsilon}.$$

It then follows immediately that

$$n \geq \frac{\alpha(1 + \epsilon) - q}{\epsilon(p + \epsilon)}$$

and, therefore,

$$\frac{T_\mu}{T_\lambda} < \left(\frac{n - \mu + \alpha}{n - \alpha + 1} \frac{p}{q} \right)^\alpha < \left(\frac{p}{p + \epsilon} \right)^\alpha \leq \eta.$$

4. *Use this bound to show the tail probability C is smaller than η .* Putting this together gives

$$C < A(\eta + \eta^2 + \eta^3 + \cdots) = \frac{A\eta}{1 - \eta} < \frac{(1 - C)\eta}{1 - \eta},$$

and so $C < \eta$.

Quod erat demonstrandum (as Bernoulli might have said).

REMARK. It is interesting to see Bernoulli uses the device of blocking terms, because this technique was later used in the twentieth century to derive limit theorems for sums of independent and identically distributed random variables.

In short, Uspensky has provided a clear and complete presentation of an interesting (and impressive) proof by Bernoulli, one that could not be found in any of the textbook literature of the time (at least in English).⁵

⁵In the 1924 edition of his book, Markov presented Bernoulli's proof but silently included improvements in it due to Nicholas Bernoulli

2.2.2 Background. There is an interesting backstory to Uspensky's presentation of Bernoulli's proof. The year 1913 was the bicentenary of the appearance of Bernoulli's *Ars conjectandi*, the famous book in which he gave his proof. To mark the centenary, Markov arranged a meeting of the St. Petersburg Mathematical Society, as well as commissioning Uspensky (who had an excellent knowledge of Latin) to write a Russian translation of Part 4 of the *Ars*; see Bernoulli (1913) and for the meeting itself, Seneta (2013). Uspensky thus had an intimate knowledge of Bernoulli's book and other material from it also appears scattered throughout Uspensky's book. (Uspensky in fact owned a copy of the *Ars conjectandi*, which is now part of the Stanford University Library Special Collections.⁶)

Good English translations of the *Ars* are now available; see particularly Sylla (2006), and also Sheynin (2005). Stigler (1986, pp. 63–70) gives an excellent historical account of Bernoulli's theorem and his proof of it.

2.2.3 Other good material in this chapter. Since our goal is to focus on the most distinctive and interesting topics in Uspensky, we are not going to go through his book, chapter-by-chapter, in a systematic way. But to give a sense of the richness of the book, here are some other highlights of just this chapter. For example, immediately after his treatment of Bernoulli's theorem, Uspensky states and proves *Cantelli's theorem*, something rarely found outside of graduate texts (and often not even then).

THEOREM 2.1. For $\epsilon < 1$, $\eta < 1$ let N be an integer satisfying the inequality

$$N > \frac{2}{\epsilon^2} \log \frac{4}{\epsilon^2 \eta} + 2.$$

The probability that the relative frequencies of an event E will differ from p by less than ϵ in the N th and all the following trials is greater than $1 - \eta$.

Other highlights of the chapter include:

- A lengthy translation from the *Ars conjectandi*.
- A survey of experimental verifications of Bernoulli's theorem.
- A discussion of the Buffon needle problem.⁷

(providing better bounds on the tail probability), as well as a “modernized” version of the proof, which dispensed with unnecessary restrictions (on n , p and ϵ). Strictly speaking, it is Markov's improved and modernized version that Uspensky gives, together with an improvement of his own (a lower bound on n that no longer depended on p). Hald (1990, Chapter 16) gives a detailed analysis of Bernoulli's original proof, as well as discussing the contributions of Nicholas Bernoulli, Markov, and Uspensky.

⁶But only recently—Steve Stigler was still able to check this out from the main library stacks in 1972.

⁷Including an extension of it to the case of a triple grid; see Perlman and Wichura (1975, pp. 159–162) for a discussion of the statistical aspect of estimating π in this case.

- The standard Chebyshev proof of the weak law of large numbers given as a three-part exercise.
- The proof of the Weierstrass approximation theorem, using Bernoulli's theorem to show the Bernstein polynomials are dense in the continuous functions on a closed and bounded interval $[a, b]$, is sketched in an exercise.

2.3 The Lexis Ratio

The middle of the nineteenth century saw an increasing interest in the apparent stability of statistical ratios (say the yearly suicide rates in Paris and Marseilles). An important issue that arose out of this was whether different sets of frequency data from different populations could be legitimately combined: that is, whether they represented trials of the same or different phenomena. Suppose one has n independent series of s independent Bernoulli trials each. Let m_1, \dots, m_n denote the number of successes in each of the n series of s trials, and let p denote the mean probability of success in all $N := ns$ trials. In 1876, the German statistician Wilhelm Lexis (1837–1914) defined the *coefficient of dispersion*

$$Q := \frac{\sum_{i=1}^n (m_i - sp)^2}{Np(1-p)}.$$

Let p_{ij} be the probability in the i th series that the j th outcome is a success. Then the mean probability in the i th series is

$$p_i := \frac{p_{i1} + p_{i2} + \dots + p_{is}}{s}$$

and the overall mean probability is

$$p = \frac{p_1 + p_2 + \dots + p_n}{n}.$$

If $D := E(Q)$, then one can show that D equals

$$1 + \frac{s-1}{np(1-p)} \sum_{i=1}^s (p - p_i)^2 - \frac{1}{Np(1-p)} \sum_{i=1}^n \sum_{j=1}^s (p_i - p_{ij})^2.$$

There are three natural special cases here:

1. The *Bernoulli* case: the p_{ij} have the same value p . In this case, $D = 1$ (“normal dispersion”).

2. The *Lexis* case: the p_{ij} are constant within strata; $p_{ij} = p_i$. In this case, the third term in D vanishes but not the second and $D > 1$ (“supernormal dispersion”).

3. The *Poisson* case: the p_{ij} vary within a stratum but in the same way from one stratum to another; $p_{i1j} = p_{i2j}$. In this case, the second term in D vanishes but not the third and $D < 1$ (“subnormal dispersion”).

There was a substantial literature available in English about this in 1937 (see, e.g., Fisher, 1922, Chapter 10; Forsyth, 1925; Rietz, 1924, Chapter 6; Coolidge, 1925, Section 4.2), but Uspensky's treatment introduced contributions by Markov and Chuprov that were unknown in the English literature of the time.⁸

Markov considered an empirical version \hat{Q} of the dispersion coefficient Q replacing the unknown theoretical probabilities by ones estimated from the data: if $M = \sum_i m_i$, then

$$\hat{Q} = \frac{n(N-1)}{(n-1)} \frac{\sum_{i=1}^n (m_i - s \frac{M}{N})^2}{M(N-M)}.$$

(If $M = 0$ or $M = N$, then $\hat{Q} = 1$ by definition.) Chuprov and Markov were then able to establish the following exact expressions for $E(\hat{Q})$ and $\text{Var}(\hat{Q})$.

THEOREM 2.2. *In the Bernoulli case $E(\hat{Q}) = 1$ and*

$$\text{Var}(\hat{Q}) = \frac{2N(N-n)}{(n-1)(N-2)(N-3)} \cdot \sum_{M=1}^{N-1} \frac{N-M-1}{N-M} \binom{M}{N} p^M (1-p)^{N-M}.$$

This immediately gives a simple bound for the variance

$$\text{Var}(\hat{Q}) < \frac{2N(N-n)}{(n-1)(N-2)(N-3)};$$

and, for $n \geq 5$ the even simpler bound

$$\text{Var}(\hat{Q}) < \frac{2}{n-1}.$$

For excellent historical accounts of the Lexis ratio and the contributions of Markov and Chuprov, see Heyde and Seneta (1977, Section 3.4), Ondar (1981) and Stigler (1986, Chapter 6).

2.3.1 Ships passing in the night. There was an interesting disconnect between the English and Continental literatures on statistics during the nineteenth and early twentieth centuries. One simple example is the distribution of S^2 , the sample variance, in the case of sampling from a normal population: Helmert had already worked this out in 1876, but this was overlooked in England and independently rediscovered by Student in his famous 1908 paper on the t -statistic (see Zabell (2008, Section 2.3.1)). Indeed, Helmert's priority was only recognized and acknowledged in the English literature much later (by Karl Pearson, 1931).

A similar situation held in the case of the Lexis ratio: although effectively the same as the chi-squared statistic, it was only in 1925 that R. A. Fisher observed:

It is of interest to note that the measure of dispersion, ϕ^2 , introduced by the German economist Lexis is, if accurately calculated, equivalent to χ^2/n of our notation [Fisher, 1925, p. 79].

Thus, much like Moliere's *bourgeois gentilhomme* who (as the Wikipedia puts it) was "surprised and delighted to learn that he has been speaking prose all his life without knowing it," both groups of statisticians had been talking about the same thing without realizing it; and one could still find papers in *Biometrika* in the 1930s and 1940s computing the exact expectation of the chi-squared statistic, in ignorance of the earlier contributions of Markov and Chuprov.

But—surprisingly—even Uspensky himself was not immune to this breakdown in communication. After discussing the exact expectation and variance of \hat{Q} , Uspensky went on to write (p. 219):

It would be greatly desirable to have a good approximate expression for the probability of either one of the inequalities

$$\hat{Q} \geq 1 + \epsilon \quad \text{or} \quad \hat{Q} \leq 1 - \epsilon.$$

But this important and difficult problem has not yet been solved.

This curiously overlooks not only the work of R. A. Fisher (with which Uspensky was in fact well acquainted), but also Markov (who had worked out the limiting distribution of \hat{Q} in 1920).

2.3.2 Uspensky on error estimation and robustness. When later discussing the closely related case of Pearson's chi-squared test for the goodness-of-fit of a vector of observed frequencies to a prescribed vector of multinomial probabilities, Uspensky argued (p. 327) that because "the degree of approximation [for the test] remains unknown," the "lack of information as to the error incurred by using an approximate expression . . . renders the application of this " χ^2 -test" devised by Pearson somewhat dubious." (This issue was later addressed in part by Cochran, 1952, among others.) Of course, similar concerns can also be raised about the use of t -statistics and the other elements of normal sampling theory, and here too Uspensky expressed (p. 345) reservations about their use:

The various distributions dealt with in this chapter are undoubtedly of great value when applied to variables, which have normal or nearly normal distribution. Whether they are always used legitimately can be doubted. At least the "onus probandi" that the "populations" which they deal with are even approximately normal rests with the statisticians.

⁸Julian Lowell Coolidge (1873–1954) was a mathematician at Harvard University who wrote a number of books on mathematics and the history of mathematics. His 1925 book was a substantial (if inferior to Uspensky's) book in English at the time on mathematical probability.

Such robustness issues had been raised earlier by Egon Pearson in his 1929 review in *Nature* of the second (1928) edition of R. A. Fisher's *Statistical Methods for Research Workers*; see [Pearson \(1990, pp. 95–101\)](#) and [Zabell \(2008, p. 4\)](#). Nowadays there is a general consensus that this is less of a concern in the case of estimates or tests for means, but much more so in the case of variances.

In general, Uspensky is typically interested not just in limit theorems as approximations, but also providing an estimate of the magnitude of the error (as in Markov's method of continued fractions).

2.4 Some Other Interesting Gems

There are many other interesting topics in Uspensky rarely found in other books, even today. These include:

- The use of difference equations in solving problems (Chapter 5).
- An analysis of the error in the normal approximation to the binomial (Chapter 7).⁹
- A bound on the error for the Poisson approximation to the binomial (Chapter 7, Exercise 9, pp. 135–136).
- Gambler's ruin with unequal stakes (Chapter 8, pp. 143–146).
- De Moivre's simple formula for the mean absolute deviation of the binomial (Chapter 9, p. 176, Exercise 1).¹⁰
- Bernstein's inequality (Chapter 10, pp. 204–205, Exercises 12–15).
- Moment inequalities (Chapter 13, p. 278, Exercise 3).
- A detailed reconstruction of Liapounov's proof of the Liapounov central limit theorem (Chapter 13, pp. 284–296).¹¹
- A central limit theorem for two-state Markov chains (Chapter 13, pp. 297–302).
- A crash course in mathematical statistics (Chapter 16).¹²

⁹[Dutka \(1981, p. 16\)](#) describes this as a “very sophisticated version of Laplace's analysis,” and notes the “considerable technical difficulties” involved in its proof. See [Seneta \(2013, pp. 1112–1114\)](#) for historical context and discussion. Interestingly, Littlewood late in life became interested in this problem, although his 1969 paper on it contains several (fixable) errors; see [McKay \(1989\)](#).

¹⁰See [Diaconis and Zabell \(1991\)](#) for the history of this most interesting result. As noted there, the formula was independently rediscovered several times, including twice after Uspensky's book appeared.

¹¹See ([Siegmund-Schultze, 2006](#)) for a critique and connections with the later work of von Mises.

¹²Uspensky gives detailed mathematical derivations of the distributions of χ^2 , S^2 , t , r and $\tanh(r)$ for samples from a normal population. This appears to be have been the first textbook in English to do so. Although all of these were of course available in the research literature, a similar exposition two years earlier in *The American Mathematical Monthly* ([Jackson, 1935](#)) justified itself on the grounds of “bringing together items that are scattered through a variety of books and journals, and in supplying explanations which in one account or another may

- Stirling's formula with bounds and the evaluation of definite integrals (Appendix 1).¹³
- The method of moments and its applications (Appendix 2).¹⁴
- Kuzmin's solution to Gauss's continued fraction problem (Appendix 3).

Several of these topics warrant brief mention because they crop up in later literature.

2.4.1 The problem of runs. One illustration of the use of difference equations Uspensky gives is to solve the *problems of runs* (Sections 3–8, pp. 77–84): to find the probability y_n of at least r consecutive successes in n tosses of a p -coin. The y_n are easily seen to satisfy the recursion

$$y_{n+1} = y_n + (1 - y_{n-r})qp^r.$$

If we let $z_n := 1 - y_n$, this gives us the recursion

$$z_{n+1} - z_n + qp^r z_{n-r} = 0,$$

and this in turn enables us to find the generating function of the z_n . Uspensky (pp. 78–79) gives the detailed argument, finding the generating function of the z_n to be

$$\phi(\xi) = \frac{1 - p^r \xi^r}{1 - \xi + qp^r \xi^{r+1}},$$

which can be expressed as a power series in ξ “according to the known rules.” The z_n are the coefficients of this power series.

Of course, for this to be useful one wants a formula for the coefficients. If one sets

$$\beta_{n,r} := \sum_{k=0}^{\lfloor \frac{n}{r+1} \rfloor} (-1)^k \binom{k}{n - kr} (qp^r)^k,$$

Uspensky says one can show “without any difficulty” that

$$z_n = \beta_{n,r} - p^r \beta_{n-r,r}.$$

For the details of this last calculation, see [Ethier \(2010, p. 41\)](#). Although the generating function for the z_n was known long before, going back to Laplace, see [Todhunter \(1865, pp. 539–541\)](#) and [Hald \(1990, pp. 417–421\)](#), Uspensky's formula for its coefficients may originate with him.

Obviously, this formula for z_n can be challenging to compute for large n , and Uspensky devotes several pages

have to be read between the lines.” But even here Uspensky played a role: in his acknowledgements at the beginning of his paper, Jackson wrote: “In preparing the paper for publication the author has derived profit from remarks made by Professors Hotelling and Uspensky in the discussion following the oral presentation.”

¹³The log convexity argument in Section 3 goes back to [Bohr and Mollerup \(1922, pp. 149–164\)](#).

¹⁴This is a difficult topic, with very few expository presentations for nonexperts. We do not know of anything comparable.

(79–84) giving approximate formulas for z_n when n is large, drawing on the results in an earlier (but uncited) paper of his (Uspensky, 1932).

The problem of runs was first solved by de Moivre (1738, Problem 88, pp. 243–248; 1756, Problem 74, pp. 254–259). De Moivre gave the following algorithm to compute the probability of a run: if $q = 1 - p$ and $c = p/q$, then expanding

$$\frac{p^r}{1 - q - cq^2 - c^2q^3 \dots - c^{r-1}q^r}$$

in powers of q and taking the sum of the first $n - r + 1$ terms “will express the probability required.”¹⁵ De Moivre did not explain where his formula came from, but Todhunter (1865, Section 325, pp. 184–186) helpfully sketched a derivation.

2.4.2 Gambler’s ruin with unequal stakes. Uspensky’s Chapter 8 (“Further Considerations on Games of Chance”) begins with a discussion of the standard topic of computing the absorption probabilities for gambler’s ruin with equal stakes, but then (p. 143) turns to the non-standard question of what happens when the stakes are unequal:

Two players A and B play a series of games, the probability of winning a single game being p for A and q for B , and each game ends with a loss for one of them. ... [If] the stakes for A and B measured in a convenient unit are α and β and their respective fortunes are a and b , find the probabilities for A and B to be ruined in the sense that at a certain stage the capital of A become less than α or that of B less than β .

If y_a is the probability of ruin for A , Uspensky (pp. 143–146) shows in the case of a fair game (i.e., $p\beta = q\alpha$) that y_a satisfies the inequalities

$$\frac{b - \beta + 1}{a + b - \beta + 1} \leq y_a \leq \frac{b}{a + b - \alpha + 1},$$

as well as a slightly more complicated set of inequalities in the unfair case.

Here, once again, Uspensky is drawing on “an ingenious method developed by Markov” (see Markov, 1900, pp. 142–146 of the 1912 edition). The problem had been studied earlier by Rouché (1888) and discussed in Bertrand (1888), the crucial expression being the equation

$$pz^{\alpha+\beta} - z^\alpha + q = 0.$$

Markov, in an obscure paper in 1903,¹⁶ showed that all the roots to this equation contribute to the probability of ruin

¹⁵As Todhunter notes, De Moivre gives an erroneous value for c in his algorithm, but the actual computations in his examples are correct.

¹⁶“On the question of the ruin of players” (K voprosu o razorenii igrokov), *Izv. fiz.-mat. obshch. pri Kaz. univ.*, 1903; see Markov (1900, p. 211 of the 4th, 1924 ed.).

and derived the inequalities Uspensky gives. Markov’s method also appears briefly in Feller (p. 366 of the 1968 edition).

Since Uspensky the problem has continued to generate a modest literature. Feller generalized Uspensky’s result in 1950 (see p. 366 of the 1968 edition of his book). Seal (1966) traces the history on this up to 1957. Later literature includes Ethier and Khoshnevisan (2002) and Gilliland, Levental and Xiao (2007), the second giving exact formulas in terms of the roots of $pz^{\alpha+\beta} - z^\alpha + q = 0$. Ethier (2010, Sections 7.2–3) discusses these generalizations at length and reviews their history (pp. 273–274).

2.4.3 Bernstein’s inequality. Let X_1, X_2, \dots, X_n be independent random variables such that $EX_j = 0$ and $\sigma_j^2 := \text{Var } X_j < \infty$ for $j = 1, \dots, n$. Bernstein’s inequality states that if $S_n := X_1 + \dots + X_n$, $B_n^2 = \text{Var } S_n$ and for some $c > 0$ one has

$$E[|X_j^k|] \leq k! \frac{\sigma_j^2}{2} c^{k-2}$$

for $j = 1, \dots, n$ and $k > 2$, then

$$P(|S_n| > t) < 2 \exp\left(-\frac{t^2}{2B_n + 2ct}\right).$$

In particular, if the X_j are uniformly bounded, that is, $|X_j| \leq M$ for some $M > 0$, then one can take $c = M/3$. This is (like a number of other interesting results in Uspensky) established via a sequence of exercises; see Uspensky (1937, pp. 204–205, Problems 12–15). Nowadays better results are available, but Bernstein’s inequality remains a simple and useful upper bound.

The inequality was proved by Sergei Natanovich Bernstein (1880–1968), and appears in his 1927 book *Theory of Probability* (as well as an earlier 1924 paper), but even in 1962 these were hard to find in the US. (Bennett, 1962, listing previous references to the inequality in the English literature, described them as being “unobtainable.”) Uspensky’s account in his book—with one exception—remained the sole source in English for the inequality for many years. Bennett (1962, p. 35), discussing the inequality a quarter of a century later, could only find six previous instances where it was mentioned in the English literature; only one of these, Craig (1933), predates Uspensky’s book or gives a derivation. But this turns out to be the exception that proves the rule. In his paper on the inequality Craig says (p. 94):

Another interesting and important attempt in this direction [i.e., Chebychev’s inequality for a sum of independent random variables] due to S. Bernstein seems to have generally escaped attention in the English-speaking world, at least, since it has been published only in Russian.

How then did Craig learn of it? He tells us in a footnote: “The present account of this work of Bernstein is taken from a lecture of Professor J. V. Uspensky”!

For other citations of Bernstein via Uspensky, see Blackwell (1954, p. 397) and Bahadur (1966, pp. 578–580).¹⁷

3. USPENSKY’S TWO OTHER TEXTBOOKS

Uspensky wrote two other textbooks during his time at Stanford, one on number theory (Uspensky and Heaslet, 1939) and one on the theory of equations (Uspensky, 1948). While our primary focus has been on Uspensky’s book on mathematical probability, his other two textbooks are not without interest, so we briefly comment on them.

3.1 Elementary Number Theory

Uspensky was first and foremost a number theorist (an interest he shared with his advisor Markov). So not surprisingly his book on number theory also contains many interesting examples and topics. Here are a few:

- Lamé’s theorem (the number of divisions needed to find the gcd of two numbers is at most five times the number of digits in the smaller number, pp. 43–45).
- Bonse’s inequality (if p_n is the n th prime, then $p_{n+1}^2 < p_1 p_2 \dots p_n$, $n \geq 4$, pp. 86–89).
- Meissel’s formula for computing the number of primes less than a given number (pp. 120–124).
- Kummer’s proof of the law of quadratic reciprocity (pp. 375–379).
- Dickson’s proof of Lagrange’s four-square theorem (pp. 379–386).
- Chapters on the Bernoulli numbers, including the Voronoi congruences (Chapter 9) and Liouville’s methods for deriving properties of arithmetical functions, culminating in an elementary proof of Gauss’s theorem characterizing integers expressible as a sum of three squares (Chapter 13).
- Appendices on magic squares (pp. 159–172), calendar problems (pp. 206–221), and card shuffling (pp. 244–248).

We cannot resist briefly discussing this last topic.

3.1.1 Card shuffling. Uspensky and Heaslet (1939, pp. 244–248) is in effect a standalone article on shuffling cards. They treat perfect shuffles and the “Monge” or “over-under” shuffle. Consider a deck of $2n$ cards (e.g., $2n = 52$) in order $1, 2, \dots, 2n$ from the top down. The

¹⁷In a note added in proof, Blackwell reports Ted Harris had drawn his attention to the Bernstein inequality, cites Uspensky, and notes that using it would yield his geometric rate of decay result under weaker conditions and with a slightly better rate. Bahadur told a colleague that he would not have been able to write his 1966 paper if he had not known of the result in Uspensky.

TABLE 1
Number of shuffles r required for a deck of size $2n$

$2n =$	2	4	6	8	10	12	14	16	18	20	22
$r =$	2	4	3	6	10	12	4	8	18	6	11
$2n =$	24	26	28	30	32	34	36	38	40	42	
$r =$	20	18	28	5	10	12	36	12	20	14	
$2n =$	44	46	48	50	52						
$r =$	12	23	21	8	52						

deck is cut exactly in half and then the two halves riffle shuffled together as follows: if the cards in order are

1, 2, 3, 4, 5, 6, 7, 8, 9, 10,

the deck is divided into two,

1, 2, 3, 4, 5 and 6, 7, 8, 9, 10,

and then interleaved:

6, 1, 7, 2, 8, 3, 9, 4, 10, 5.

This is a perfect “in shuffle,” practiced by gamblers and found in magazines since (at least) 1743.

The question is, how many perfect shuffles (r) are required to recycle the cards—that is, to bring them back to their original order? When one of us was quite young (PD), we figured this out “the long way,” by actually shuffling the cards until they returned to their original order. Some of the data is given in Table 1.

In their book, Uspensky and Heaslet prove that after a single shuffle card i moves to position $2i \pmod{2n+1}$, so after k shuffles a card in position i moves to position $2^k i \pmod{2n+1}$. It follows the deck returns to its original order the first time $2^k \equiv 1 \pmod{2n+1}$. Suppose $2n+1$ is prime (e.g., when $2n = 52$). Observe that in this case $2n$ shuffles are *sometimes* required before the deck returns to its original order. Does this happen infinitely often? Nobody knows—this is a special case of the *Artin conjecture* (that 2 is a primitive root for infinitely many p). The conjecture is known to be true on the generalized Riemann hypothesis but this is a million dollar problem.

As far as we know, Uspensky and Heaslet were the first to discover and prove that the number of shuffles required to return the deck to its original configuration is the order of two. For the history of perfect shuffles and their applications, see Diaconis, Graham and Kantor (1983), Diaconis and Graham (2012). Uspensky and Heaslet’s Appendix (pp. 245–248) also treats the “Monge shuffle” with similar results. This particular shuffle was first analyzed by the French mathematician, Gaspard Monge, in 1783.

Chapter 13 (“Liouville’s Methods”) was a bit of an indulgence on Uspensky’s part, because it discussed a major research interest of his, the results being “elementary” only in terms of the methods used. The chapter derived

a variety of very general arithmetical identities using basic number theory arguments rather than appealing to the theory of elliptical functions. Uspensky then used special cases of these identities to derive the Gauss and Jacobi theorems on the representation of an integer as a sum of three or four squares.

Of course, Uspensky's book did not touch on most of his research (it was, after all, on elementary number theory), but before leaving it, we discuss one other number theoretic result of Uspensky's, one not treated in his book but which remains surprisingly fresh even today.

3.1.2 A game. For α a positive real number, define the *spectrum* of α to be the sequence of integers

$$\alpha_n := \{ \lfloor n\alpha \rfloor : n = 1, 2, 3, \dots \},$$

where $\lfloor x \rfloor$ is the floor of x . Such sequences are termed *Beatty sequences* (named after Samuel Beatty, 1881–1970). A classical theorem attributed to Beatty says that if $\alpha, \beta > 1$ are a pair of irrational numbers such that $1/\alpha + 1/\beta = 1$, then the spectra of α and β are disjoint and cover the natural numbers. (This appears on p. 98 of Uspensky and Heaslet as Exercise 9.) It is natural to ask about a similar triple α, β, γ (or more). Uspensky (1927) proved coverage with disjoint spectra only occurs in the case of two spectra. His proof is involved, using Kronecker's approximation theorem. An elementary proof due to Graham (1963) led to fascinating further theory and many still open problems; see Graham, Lin and Lin (1978), Graham and O'Bryant (2005).

Curiously, even before Beatty the result was stated by Lord Rayleigh (John William Strutt). Lord Rayleigh (Strutt, 1894, p. 123) says:

Thus, if x be an incommensurable number less than unity, one of the series of quantities $m/x, m/(1-x)$, where m is a whole number, can be found which shall lie between any given consecutive integers, and but one such quantity can be found.

Rayleigh did not provide a proof. Beatty (1926) posed it as a problem in the *American Mathematical Monthly*, and solutions were submitted by several contributors (Beatty et al., 1927). Apparently independently, Willem Abraham Wythoff (1865–1939) proposed in 1907 a modification of the game Nim in which the “cold” (losing) positions are parametrized by the pair of complementary Beatty sequences generated by ϕ and ϕ^2 , where ϕ is the golden ratio $(1 + \sqrt{5})/2$. (Uspensky discusses Nim at some length in his book, pp. 16–19, but does not note this connection.)

3.2 The Theory of Equations

Uspensky's last book, *Theory of Equations*, appeared posthumously. Sent to the publisher in December 1946, only a month before he died (on January 27, 1947), the

manuscript was seen through the press by Uspensky's former students Max Heaslet and Carl Olds, and appeared in 1948. It is less distinctive than the other books: it was, as Uspensky notes in the preface, “elementary in nature and, with few exceptions, contains only material customarily included in texts of this kind.” It was longer than other texts on the subject then currently in use, being designed for self-study if desired.

But—Uspensky being Uspensky—there were still elements of novelty, ones in which “the exposition differs considerably from custom.” In Chapter 1, on complex numbers, “the superficial approach so common in many books” was replaced by the rigorous definition of a complex number as an ordered pair (a, b) of real numbers. In Chapter 6, a method for separating real roots was given based on Vincent's theorem.¹⁸ In Chapter 7, on the approximate evaluation of roots that had been separated, Sections 2–7 were devoted to the original form of Horner's method, “which unfortunately has disappeared from American texts.” In Chapter 9, determinants were introduced using Weierstrass's approach, based on their “characteristic properties” rather than “formal definition.” (This is similar to Artin's axiomatic development of them, which can be found in some editions of Lang's calculus and linear algebra textbooks.) And in an Appendix, Uspensky gave Gauss's fourth proof of the fundamental theorem of algebra.

All of Uspensky's books are marked by an exemplary intellectual sharpness: the presentation is easy to follow, the proofs are complete, readable and coherent, in some cases the results are given in greater generality than is ordinarily the case. The writing, thinking through of the material, order of presentation and choice of examples is both engaging and natural.¹⁹ The references given to the prior literature are invariably to the masters, not the

¹⁸Vincent's theorem, named after Alexandre Vincent (1797–1868) and published by him in 1834, appears to have been entirely forgotten until Uspensky's discussion of it in his book (see Chapter 6, Section 12 and Appendix 2). Uspensky describes the method as “very efficient” and “much superior in practice to that based on Sturm's theorem,” adding he “believes that no other book mentions this method, which he invented many years ago and has been teaching to his students for a number of years” (p. v). Uspensky's Appendix 2 gave a sharper version of Vincent's result, providing a bound on the number of steps in the algorithm. Shortly after, Ostrowski (1950) was able to improve Uspensky's bound (in a paper which was itself however overlooked for many years). Ostrowski notes in passing (p. 702, footnote 4) a typographical error in one of Uspensky's conditions, where a 2 appears instead of a 1/2, something a later, less perceptive reader mistook for an actual error in Uspensky's proof. For later literature, see Krandick and Mehlhorn (2006).

¹⁹“As a teacher, he followed the classical style and ideals. His presentation, whether oral or in writing, was clear, simple, logical and elegant. He was able to present the simplest ideas in a fascinating way” (Pólya, Szegő and Young, 1947, p. 2).

pupils; books and papers are not just cited for show but are ones Uspensky has both read and mastered.

Uspensky grew up in the classical Russian tradition; and—perhaps as a result—he exhibits an impressively broad mathematical culture. This suggests some natural questions: who was he, what was his mathematical background, how did he come to the US?

4. A CLOSER LOOK AT USPENSKY

Detailed information about Uspensky is not easy to find in English, so in this section we discuss his life, based in part on family records, unpublished material in the Stanford archives, material available only in Russian and many passing references to him throughout the English literature.

4.1 Uspensky in Russia

Yakov Viktorovich Uspenskii was born on April 29, 1883, the fourth of five children, in Urga, Mongolia.²⁰ (His father Viktor Matveevich Uspenskii, 1845–1901,—a career diplomat—was the Russian consul there.²¹) When Uspensky was seven, his mother and her children moved to St. Petersburg to enable the children to receive a proper Western education, while his father (then in the Chinese province of Sinkiang where he was Russian consul-general for Western China) remained at his diplomatic post except when on vacation in Russia.

Uspensky attended a classical *gymnasium* in St. Petersburg; it was here he learned both Greek and Latin, knowledge which was to prove useful later in his historical studies. Although initially interested in astronomy, early on he became interested in mathematics; by the time he graduated, he had an excellent grounding in the differential and integral calculus, and was able to read books on astronomy and theoretical physics. He graduated from the *gymnasium* in 1902 with distinction, receiving a gold medal in recognition of his achievements.

4.1.1 The Russian university system. In order to understand Uspensky's university career, a few words about the Russian system then in place may be helpful. Starting at the beginning of the nineteenth century, there were three academic degrees: the *kandidat*, *magister* and *doktorat*. The *kandidat* (later the *diplom*) was a bachelor's degree,

awarded after completing one's studies and passing a set of examinations; the *magister* and *doktorat* were graduate degrees, awarded after writing and publishing a thesis, which then had to be defended in a public oral examination (in the case of the *magister*, one also had to take and pass a set of examinations prior to the writing of the thesis); see generally Sanders (1993).

The *magister* (Latin for “master” or “teacher,” as in *Magister Ludi*) was the equivalent of today's Ph.D., and required in order to teach as a *privat dozent*, that is, an instructor who took private pay students.²² The *doktorat* was a second doctoral decree, required the writing of a second thesis, and was usually necessary in order to become a professor at a university. There were finally in turn two grades of professorship, termed “extraordinary” and “ordinary,” roughly equivalent to being an Associate and Full Professor, respectively.²³

4.1.2 Uspensky's university career. Uspensky spent his entire academic career in Russia at the university in what was then St. Petersburg.²⁴ He was an undergraduate there from 1903–1906, years which overlapped the first Russian revolution of 1905 and saw a considerable disruption of Russian academic life. But unlike many of his fellow students he did not become involved in politics and devoted himself entirely to his studies. (His apolitical nature may in part explain his escaping relatively unscathed during the upheavals of the 1920s.) Even at this early stage his talent was evident: while still an undergraduate he wrote his first paper, the first rigorous proof that the cyclotomic ring $\mathbb{Z}[\zeta_5]$ (ζ_5 a primitive 5th root of unity) was Euclidean (Uspensky, 1906); this work was regarded as sufficiently important that a reworked version

²²The *magister* is sometimes referred to as a master's degree, but this is misleading given the current academic status of the master's degree relative to the doctorate. In contrast, in the medieval universities of Europe the titles Master and Doctor were effectively synonymous, and which degree was conferred depended on both the university and the faculty within it (arts, law, theology, medicine); see Rashdall (1895, pp. 21–22) and Verger (2003, p. 146).

²³The terminology derives from the Latin words *extraordinarius* and *ordinarius*, where the “extra” connotes difference or separateness (as in extralegal or extraterrestrial) rather than superiority. The system was similar to the German one at that time, where the two ranks were *außerordentlicher* and *ordentlicher Professor*, professorships without and with a chair, respectively.

²⁴The university underwent several name changes during the period we will be discussing: Saint Petersburg Imperial University (1821–1914), Petrograd Imperial University (1914–18), Petrograd State University (1918–24), and Leningrad State University (1924–91). The city of St. Petersburg was renamed Petrograd immediately after the outbreak of war because both the “Sankt” and the “burg” in “Sankt-Peterburg” were German words. The second name change followed the overthrow of the Czarist regime in 1917, and the third reflected the change of the name of the city from Petrograd to Leningrad five days after Lenin died on January 21, 1924.

²⁰Uspensky almost invariably signed his books and papers in English as “J. V. Uspensky.” The sole exception is his *Annals of Mathematics* paper of 1927, in which he uses the name “James V. Uspensky.” His Declaration of Intention for naturalization, dated March 24, 1930, gives his full name as “James Viktorovitch Uspensky”; but his 1942 Registration Card for the draft shortened this to “James Victor Uspensky.”

²¹See ru.wikipedia.org/wiki/Успенский,_Виктор_Матвеевич (accessed on March 8, 2020) for further information about Uspensky's father.

of it written in French was published in the *Mathematische Annalen* in 1909. In 1906, Uspensky graduated, receiving a *diplom* of the first degree (i.e., with honors), was awarded a scholarship, and immediately began his graduate studies. By 1908, he completed his course work, passed his examinations and began working on his thesis under the direction of Markov. The thesis (“Some applications of continuous parameters to number theory”) was completed, approved and published in 1910; after passing his oral examination the next year in 1911, Uspensky was then formally awarded the degree of *Magister of Pure Mathematics*.²⁵

With his *magister* in hand, Uspensky became a privatdozent at St. Petersburg in 1912 (although he also supplemented his income during this period by teaching at other institutions: the Institute of Railway Engineers, 1907–29, the Higher Women’s Courses, 1911–17). In 1915, he became an extraordinary professor at Petrograd Imperial University; this was unusual because he had not yet been awarded his *doktorat*, but presumably reflected both his ability and a wartime shortage of personnel. He was in turn rapidly promoted to ordinary professor in 1917; this again reflected his ability but the abolition of the *doktorat* around this time may also have played some role as well.

In 1921, Uspensky was elected to the Russian Academy of Sciences (as a replacement for Liapunov, who had died in 1918), his election supported by Markov, Steklov and Krylov (1921). Their report, besides giving a complete bibliography up to that time, cited several of his papers as justifying this honor, in particular (Uspensky, 1920), a paper in which Uspensky had derived (independently of Hardy and Ramanujan, 1918) the asymptotic formula for the number of partitions of an integer,

$$p(n) \sim \frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2}{3}n}\right),$$

as well as an estimate for the error.²⁶ (Because of wartime and revolutionary conditions, the Hardy and Ramanujan paper was unknown in Russia, even in 1921.)

²⁵Most sources say Uspensky received his Ph.D. or doctorate in 1910; although essentially correct, the actual title and year the degree was awarded are as indicated.

²⁶Although Uspensky did not derive the full divergent series expansion for $p(n)$ that Hardy and Ramanujan found using their celebrated “circle method,” he arrived at the simpler formula displayed above by deriving the first term in the expansion together with an estimate of the error, namely

$$p(n) = \frac{e^{\pi\sqrt{\frac{2}{3}(n-\frac{1}{24})}}}{4\sqrt{3}(n-\frac{1}{24})} \left(1 - \frac{\sqrt{3}}{\pi\sqrt{2n-\frac{1}{12}}}\right) + \rho_n e^{\frac{\pi}{\sqrt{6}}\sqrt{n}},$$

where the ρ_n are bounded (Uspensky, 1920, p. 209, but correcting a typographical error in Uspensky’s formula where a $\sqrt{\pi}$ appears instead of a π). This corresponds to Hardy and Ramanujan’s equation (1.55) on page 82. As they noted, this “is an asymptotic formula of a type far more precise” than the simpler one, and it was “with considerable surprise that we found what exceedingly good results the formula gives for

4.2 From Petrograd to Palo Alto

Thus by 1921 Uspensky was a Professor at the Petrograd State University and a member of the Russian Academy of Sciences, an established and senior member of the Russian mathematical community. And yet eight years later he chose to emigrate to the US. It is natural to suppose that this was connected with the Russian Revolution—and this was indeed the case—but why did it take so long? There are some complexities here.

In 1918, *Sovnarkom* (the Council of People’s Commissars, that is, the newly established Bolshevik government) began to pass decrees designed to ensure new educational opportunities for workers and peasants. These measures were resisted by some university faculty concerned that unqualified students might be admitted. Uspensky, for example, wrote that

[Since] to succeed at the university, a student should be adequately trained, prospective students must be admitted at the university in virtue of their knowledge, not their class affiliation or political commitment.²⁷

The response of the Bolshevik government was predictable: measures were swiftly enacted to “reeducate the bourgeois professors” as well as punitive measures such as preventive detention, exile and even execution (Nazarov and Sinkevich, 2019, p. 50).

These measures were not confined to university professors. In the summer of 1922, a list of 217 perceived anti-Soviet intellectuals (writers, professors, scientists, etc.) was drawn up and the GPU (*Gosudarstvennoye politicheskoye upravlenie*, the State Political Directorate) proceeded to arrest them beginning on the night of August 16–17.²⁸ Many of those detained were then exiled; at least 160 of these by ship from Petrograd to Stettin (today the Polish city of Szczecin) on the German ships *Oberbürgermeister Hacken* on September 2 and *Preussen* on November 15. Despite his initially voicing opposition to Soviet university reorganization, Uspensky was not included among these, perhaps in part because he was perceived as being apolitical, perhaps in part because of his abstruse field (as opposed to history, literature, politics and philosophy), and perhaps because of the prestige accruing from his membership in the Russian Academy of Sciences. But

fairly large values of n .” (Uspensky’s bound on the error is less precise than Hardy and Ramanujan’s, however, since in fact $\rho_n = O(1/n)$; see generally DeSalvo, 2021.)

²⁷Central State Historical Archives of St. Petersburg. F. 7240, Schedule 14, No. 16, L. 185 recto; cited in Nazarov and Sinkevich (2019, p. 6).

²⁸The GPU was the successor of the *Cheka* (the first Soviet secret police, headed by the sinister Felix Edmundovich Dzerzhinsky, 1877–1926), and combined both internal security and foreign intelligence functions.

a chilling message had been sent as to the possible limits of dissent.²⁹

In 1924, Uspensky attended the International Congress of Mathematics in Toronto and spoke there about work with his student, Boris Venkov (Uspensky and Venkov, 1928); This may have been Uspensky's first trip outside Russia, and he took advantage of the opportunity by afterwards visiting the University of Chicago and the University of Michigan, lecturing on Russian contributions to number theory before returning home. Either because he felt after this trip that publishing his work almost exclusively in Russian had limited its impact in the West, or because he had already begun to contemplate emigrating to the US, from this point on there was a marked shift in how and where Uspensky chose to publish his work. Although he continued to submit papers to the *Bulletin of the Academy of Sciences of the USSR* during the next two years (1925–1926), all but one of the nine—a memorial notice about Steklov primarily of interest to his Russian readers—now appeared in French.³⁰

Uspensky seems to have liked what he saw in the United States, because he returned two years later, teaching at Carleton College in Minnesota during the academic year 1926–1927, followed by short lecture courses at Berkeley and Stanford. His time at Carleton marks the point at which he first started to publish in English and in US journals (fifteen papers between 1927 and 1935). During this trip, he also met his future wife, Lucile Zander, who was working in the Carleton publicity department; they married on October 13, 1927, and left for Russia shortly after.

A chat with the OGPU. When he returned to what was now Leningrad, Uspensky was interrogated by the OGPU (the Joint State Political Directorate, the successor to the GPU after the formation of the USSR in late 1922).³¹

²⁹For further information about the 1922 expulsions, see Finkel (2003) and Chamberlain (2006). Inevitably, some of the younger Russian mathematicians began to leave of their own accord, including three of Uspensky's Petrograd/Leningrad colleagues: James Alexander Shohat (1886–1944), in 1923, and Jacob David Tarmarkin (1888–1945) and Abram Samoilovitch Besicovitch (1891–1970) in 1924/1925. It is possible this may have had some influence on Uspensky's own decision. It is plausible Uspensky kept in contact with one or more of these after he emigrated (on p. 160 of his *Introduction to Mathematical Probability* he thanks Besicovitch for a proof), but to precisely what extent is a matter for conjecture.

³⁰The Russian Academy of Sciences became the Academy of Sciences of the USSR in 1925, and the name of its *Bulletin* changed accordingly.

³¹Royden, who reports this incident, identifies Uspensky's interrogator as a member of the NKVD (the People's Commissariat for Internal Affairs), saying it was a predecessor of the KGB, but this is an anachronism: the NKVD only acquired the secret police and foreign intelligence functions of the OGPU in 1934. (In general, the study of the Soviet security apparatus is complicated by its numerous reorgani-

[He was asked] how he liked America. Uspensky disarmed his interviewer by saying, "I loved it. It is a place of great opportunity, and if only I were a young man I would emigrate. But I am a member of the Academy of Science, and my career is established here. I am too old to start over again." The NKVD agent evidently reported that Uspensky was reliable and sound in his views. Thus, when Uspensky did decide to come to America a few years later, he came in style on a Soviet ship with his passage paid for by the government. [Royden (1989, p. 243)]

(Uspensky's response would have been most ill-advised a decade later at the height of the Great Terror.)

But given his entirely plausible answer to the OGPU—for Uspensky certainly was an established figure in Russian mathematics—why did he leave Russia just a year and a half later? There are said to have been a number of contributing factors. The Russian historian, Natalia Ermolaeva (1997), reports:

Returning to the USSR, Uspensky resumed his numerous duties. However, in the summer of 1929 when he again went on a business trip to the USA, he did not return to his homeland. The decision to emigrate was caused by various reasons. One of them was that during his second trip to America, U. married, and his wife categorically refused to live in the USSR. At the same time, the situation in the mathematical life of the country deteriorated sharply—there was an intensive introduction of Marxism into mathematics accompanied by persecution of scientists, including Nikolai Maximovich Günter [a close colleague of Uspensky].

The final straw appears to have been a tumultuous meeting of the Leningrad Mathematical Society (LMO): G. G. Lorentz (who was an eyewitness) reports that

Arriving at the lecture, I found the entire mathematical Leningrad present. Uspensky, on the podium, was pointed out to me. After the lecture, Günter, presiding, invited the audience to pose questions. Suddenly Leifert [a Bolshevik, himself later a victim of the Great Terror] climbed on the podium yelling insults at the LMO and Günter. Many students applauded Leifert and shouted. The meeting was

zations and consequent name changes that took place during the period from 1917 to 1995: the Cheka, GPU, OGPU, NKVD, NKGB, MVD, MGB, KGB, and (after the fall of the Soviet Union) the FSB and SVR.)

dissolved. In the wake of this disastrous event, using their still valid visas, Uspensky and his wife left for the United States.³²

Or maybe not; perhaps Uspensky had in fact already made up his mind. An obituary for Uspensky in *The Sheboygan Press* (in Wisconsin, where Uspensky's father-in-law Otto Zander had been the editor) states:

When it came time for Mr. Uspensky to return to Russia [after his year at Carleton], with his exchange period expired, his wife returned there with him and they lived for a year in Russia, during which time he *completed* arrangements to return to America, leaving his native land to take up citizenship in the new country he had adopted. [February 3, 1947, p. 8; emphasis added.]

If this is accurate—and inasmuch as the information in the obituary clearly came from the family there is no reason to doubt it—Uspensky had already decided to emigrate to the US *before* he returned to Russia. (Of course, the deteriorating conditions he encountered there on his return, as Stalin increased his grip on the country, would not have helped.) Indeed, one can be forgiven for conjecturing that Uspensky had even earlier still, when he arranged to visit Carleton—a relatively obscure college from the perspective of a distinguished member of the USSR Academy of Sciences—already contemplated leaving Russia for the stability of the North American continent he had experienced during his visit to the ICM in 1924.

There is one additional item of evidence one can advance to support this hypothesis. In the first of his two papers in the *Transactions* of the AMS for 1928 (on the representation of numbers by quadratic forms), referring to his earlier work on this subject in a series of ten papers in Russian journals that had appeared between 1913 and 1926, Uspensky says that he had published these earlier investigations “*so far as it was possible under the circumstances.*” We take this to be an indirect but clear indication that when Uspensky left for Carleton he had already found the working conditions in revolutionary Russia to have significantly interfered with his ability to do mathematics.

³²Lorentz (2002, p. 189). Lorentz goes on to add: “where he accepted a Stanford University professorship offered to him by Szegő.” Here, Lorentz errs: Szegő only left Germany for the US in 1934 and only arrived at Stanford in 1938. (Szegő may of course have played some role from afar in facilitating Uspensky's appointment.)

According to the passenger manifest, Uspensky left Hamburg for the US on April 30, 1929, aboard the Hamburg–Amerika Linie transatlantic steam oceanliner *Reliance*, which arrived in New York via Southampton, Cherbourg and Cobh (Ireland). His wife Lucile seems to have left earlier, because she is not listed on the manifest. Contrary to what Royden (1989, p. 243) suggests, Uspensky did not come on a Soviet ship and traveled second class.

4.3 Stanford

Uspensky spent the summer of 1929 teaching at Carleton,³³ and then moved to Stanford that Fall. Initially appointed Acting Professor of Mathematics for the first two years (1929–1931), he became a permanent member of the faculty in 1931. This presumably came about thanks to the efforts of the chair of the Department of Mathematics, Hans Frederik Blichfeldt (January 9, 1873–November 16, 1945).

Blichfeldt, who was born in Denmark but came to the US in 1888 when his family emigrated, had impeccable mathematical credentials: after graduating with an AB from Stanford in 1896, he—like many US mathematicians at the time—went to Europe for his doctorate, studying at the University of Leipzig under the great Sophus Lie and receiving his Ph.D. there in 1898.³⁴ He then returned to Stanford, where he remained for the rest of his professional life, becoming a Full Professor in 1913, a member of the National Academy of Sciences in 1920, and Department Chair from 1927 to 1938. He did important work in group theory and number theory (including coauthoring a book with Dickson). For further information about Blichfeldt, see his obituary in the *Bulletin* of the AMS by Dickson (1947), his National Academy of Sciences biographical memoir by Eric Temple Bell (1951), his entry in the *Dictionary of Scientific Biography* (Miller, 1970) and Royden's history of the Stanford Mathematics Department (Royden, 1989, pp. 238–9).

Royden (p. 244) relates that after Blichfeldt became chair in 1927, “the Stanford mathematics department had a steady stream of major mathematicians as visiting faculty, mostly for the summer quarter.” (The Europeans among these included Harald Bohr, Edmund Landau and Gabor Szegő.) It was thus not surprising that he seized the opportunity of hiring Uspensky when the latter moved to the US. This was a major commitment at the time because the Stanford Mathematics Department was then quite small: Royden (p. 248) reports that in 1938, the year Blichfeldt retired, the department had only three Full Professors: Blichfeldt, Uspensky and W. A. Manning.³⁵ (Szegő joined Stanford in 1938, Pólya in 1942.)

Up until 1935, Uspensky published a steady stream of research papers, but then abruptly changed the form and

³³Lecturing on number theory in the first term (June 18–July 27) and recent advances in mathematical probability in the second (July 29–August 31), as well supervising reading in advanced mathematics during both; see Notes and News, *The American Mathematical Monthly* 36 (3), p. 421.

³⁴The Mathematics Genealogy Project gives the year of Blichfeldt's doctorate as 1900, but this was actually the year his thesis was published in the *American Journal of Mathematics* (Blichfeldt, 1900).

³⁵William Albert Manning (1876–1972) was Stanford's first Ph.D. in Mathematics (1904), taught at Stanford for 40 years, and was the father of Laurence Albert Manning (1923–2015), himself later a professor in Stanford's Department of Electrical Engineering for 40 years.

direction of his work in several ways. First, he turned to writing books instead of papers, his books on probability, number theory and the theory of equations appearing in 1937, 1939 and 1948 (the last finished in late 1946, shortly before his death). At the same time, he developed an interest in applied mathematics: in the last decade of his life “he took an active part in Applied Mechanics seminars and presented there many talks of great interest” (Pólya, Szegő and Young, 1947, p. 2). His colleague and fellow refugee, the engineer Timoshenko, who would have participated in these, recounted shortly after Uspensky’s death that

He participated in a seminar on applied mechanics and gave talks on various topics of applied mathematics, in such a way as to show clearly to engineers the importance of mathematics in treating engineering problems. The students appreciated these talks very much, and enjoyed also the informal discussions after seminar in which J. V. usually participated by telling them stories and anecdotes from the biographies and memoirs of famous scientists, about whom he had a great fund of knowledge. [Timoshenko (1947, p. 6)]³⁶

In the last three years of his life, Uspensky also “undertook the study of Spanish” (Pólya, Szegő and Young, 1947, p. 2) and published five papers in Spanish during this period.

Toward the end of 1946, Uspensky suffered from ill health and, although continuing to write and lecture, planned to retire from Stanford at the end of the 1946–47 academic year. But he suffered a heart attack on Thursday, January 23, 1947, and was taken to the Palo Alto Hospital, where he died the next Monday at 2:30 in the afternoon.³⁷

APPENDIX A: USPENSKY’S STUDENTS

A.1 Students in Russia

Despite his wide range of mathematical interests, Uspensky was first and foremost a number theorist. Of the five leading number theorists in Leningrad in the 1920s who did not leave Russia—Delone, Ivanov, Kuzmin,

Venkov and Vinogradov (Demidov, 2015, p. 89)—three were students of Uspensky.³⁸

Ivan Matveevich Vinogradov (September 14, 1891–March 20, 1983) was both Uspensky’s first and best-known student. Vinogradov was a central figure in modern analytic number theory, a member of the USSR Academy of Sciences, the Director of the Steklov Institute of Mathematics in Moscow for half a century (from 1934 until his death), and a Fellow of the Royal Society of London. His best known result is that every sufficiently large odd number is a sum of three primes. His *Foundations of the Theory of Numbers*, first published in 1936, is still in print today in an English translation of the 5th edition. He is too well known to require discussion here; see his Royal Society biographical memoir (Cassels and Vaughan, 1985) and Karatsuba (1981).³⁹

We have already encountered *Rodion Osievich Kuzmin* (October 9, 1891–March 24, 1949): his 1928 solution of Gauss’s challenge problem to Laplace was given in the third Appendix of Uspensky’s *Introduction to Mathematical Probability*. After completing his undergraduate studies at Petrograd State University in 1916, Kuzmin continued on to do graduate work there, but was forced to leave in 1918 due to wartime conditions. It was during this short period of time that he would have studied under Uspensky. Like Vinogradov, Kuzmin moved to the University of Perm, where he stayed until 1922, returning then to Petrograd and its Polytechnic Institute, where he remained for the rest of his life. In 1935, he received his *doktorat*, shortly after this degree was reinstated by the Soviet government.⁴⁰ In addition to Gauss’s problem,

³⁸The course of graduate study in Russia during the 1918–1934 period is complicated. On October 1, 1918, the Council of People’s Commissars abolished all academic ranks and advanced degrees. Instead, advanced training was initially supported by “professorial scholarships” for students preparing for an academic career. After 1925, the process of selection and mentoring for such studies was formalized and participants termed “aspirants” (*aspirantov*). By a decree of September 19, 1932, the standards for such programs were tightened and in research institutes the writing of a dissertation became required after a two- to three-year program of study. Finally, on January 13, 1934, academic ranks and advanced degrees were reinstated; see (DeWitt, 1961, p. 422), on which the information in this paragraph is based.

³⁹The Mathematical Genealogy Project entry for Vinogradov, while listing Uspensky as his advisor, does not list a degree, year when it was awarded or title of the dissertation. Presumably this was because of the abolition of degrees noted in the preceding footnote.

⁴⁰The Mathematics Genealogy Project states Kuzmin received the degree of Doctor of Sciences in 1935 and identifies his advisor as Uspensky, but the actual circumstances seem unclear: Kuzmin had already been teaching for many years in 1935 and Uspensky had left Russia six years earlier. The 1935 degree may have been awarded on the basis of a dissertation drawing on prior work but only submitted after the reinstatement of advanced degrees, but even so it seems highly improbable that at the onset of the Great Terror an émigré who had resigned from the USSR Academy of Sciences would have been listed as Kuzmin’s advisor.

³⁶Thus an unpublished remembrance written shortly after Uspensky’s death. Such talks were not always successful however from Timoshenko’s perspective: writing many years later (Timoshenko, 1968, p. 337), he remembered that when he asked Uspensky to lecture on partial differential equations, the lectures were “purely theoretical” and not what the engineering students needed.

³⁷Some sources give the place of death as San Francisco, but the local obituaries of the time all state he died at the Palo Alto Hospital. The details of when he was stricken, and when and where he died, have been drawn from these.

Kuzmin proved in 1930 that if a is algebraic and b is a real quadratic irrational, then a^b is transcendental. This answered a then open problem, the nature of the number $2^{\sqrt{2}}$ (the so-called Gelfond–Schneider constant). This was a special case of the celebrated Gelfond–Schneider theorem (proved in 1937), which assumes only that b is irrational. Other notable work include his 1927 paper on the *Kuzmin–Landau inequality*. For further information see, [Venkov and Natanson \(1949\)](#), [Sviderskaya \(2002\)](#).

Boris Alekseevich Venkov (July 31, 1900–December 13, 1962) worked with Uspensky while still a student; their work was later reported on in a joint paper presented to the International Congress of Mathematicians in 1924. Venkov graduated from the Leningrad State University in 1925 and became a professor there in 1935; he worked in number theory. His contributions include an elementary proof in 1928 of the Dirichlet formulas for several of the classes of binary squared forms ([Wenkov, 1931](#) is a later version in German), a theory of reduction for positive-definite quadratic forms (1940) and a characterization of polyhedra (1954 and 1959). *Venkov graphs* are named after him. For further information, see [Malyshev and Faddeev \(1961\)](#).

A.2 Students at Stanford

Uspensky supervised the doctoral dissertations of five students at Stanford. The first of these, *Harold Maile Bacon* (January 13, 1907–August 22, 1992), received his Ph.D. from Stanford in 1933; his thesis (“An Extension of Kronecker’s Theorem”) on a topic in number theory. He spent his entire career at Stanford, primarily as a long-time instructor of calculus courses, and was known to generations of Stanford undergraduates.

Bacon was a “hand-off” to Uspensky from Harald Bohr. Years later Bacon related the following charming story as to how this came to be:

Bohr was a very kind man. I remember my being in Professor Blichfeldt’s office shortly after I returned to Stanford in 1930 to continue my graduate work after my master’s degree and a year’s absence working for an insurance company under the mistaken impression that I wanted to become an actuary. Blichfeldt and I were discussing my getting started on work that might lead to a dissertation. Just then Bohr came into the office. Blichfeldt turned to him and, indicating me, said “Here’s a man who is looking for a thesis topic. How would you like to suggest one, and be his adviser?” Bohr bowed, smiled and very courteously replied, “I should be honored.” He generously acted as my supervisor for the remainder of the year he was at Stanford. When he left, I was most fortunate to have Uspensky take over and see

me through to the completion of my work on the dissertation. It was indeed a great privilege to have two such inspiring men as my friends and advisers at the beginning of my career. [[Royden \(1989, pp. 245–246\)](#)]

Because Bacon was such a visible figure at Stanford, a considerable amount has been written about him; see [Royden \(1989, pp. 244–247\)](#), [Royden et al. \(1992\)](#), [Stanford \(1992\)](#), [Jellison \(1997\)](#) and [Alexanderson \(2011, Chapter 3\)](#).

Maxwell Alfred Heaslet (February 17, 1907–July 13, 1976), Uspensky’s next student at Stanford, received his Ph.D. there in 1934 (dissertation “Concerning the Development Coefficients of an Aequianharmonic Function”). Information about him is surprisingly scanty. Here is what we have been able to find, piecing together several scattered sources.

Born in Bentonville, Arkansas, the son of Walter Monroe Heaslet and Nancy Angeline Austin, Heaslet received both his bachelor’s and master’s degrees from the University of Oklahoma before moving to Stanford in 1928. After marrying Helen Virginia Camp in June 1935, he joined the faculty of San Jose State College (now University), where he taught mathematics and physics for seven years. He began working at the Ames Research Center (housed in the Moffett Federal Airfield in nearby Mountain View) in 1942. He appears to have been initially on loan from San Jose, because he is listed in a departmental history ([Jackson, no date](#)) as being a member of the faculty from 1935 to 1945. If so, this arrangement presumably ceased at the end of the war, and Heaslet elected to stay on at Ames; he eventually became head of the theoretical aerodynamics branch of the Theoretical and Applied Research Division. He often collaborated with Harvard Lomax (1922–1999), who had a distinguished career in aeronautics and computational fluid dynamics ([Seabass, 2002](#)), and wrote many papers during his time at Ames; see [Hartman \(1970\)](#). Due to ill health, Heaslet was forced to retire from Ames in 1959 and moved to Florida where he remained for the rest of his life. He had two children, Austin and Jonathan Heaslet.

From a purely academic standpoint, Uspensky’s most successful student was *Orville Goodwin Harrold, Jr.* (September 2, 1909–May 16, 1988); he had 14 doctoral students and 135 “descendants,” and at different stages in his career was active in research (writing dozens of papers published in leading journals, including one each with Eilenberg and Moise), university administration, and professional service. He received all of his degrees from Stanford (AB 1931, AM 1932, Ph.D. 1936); although his dissertation was on a topic in numerical analysis (“On the Expansion of the Remainder in the Open Type Newton–Cotes Quadratic Formula”), he switched fields to topology immediately after he received his doctorate. The

course of his initial career presumably reflected the job market of the time: during the next decade, while holding a succession of short-term positions (Oregon State College, 1937–39, a National Research Fellow at Virginia and Michigan, 1939–40, Northwestern, 1940–42, Louisiana, 1942–43, Pomona, 1943–46, Princeton, 1946–47), he published at least ten papers before being appointed a Professor at the University of Tennessee at Knoxville, where he taught from 1947 to 1964. In 1964, he moved to Florida State University, where he was the head of the Department of Mathematics from 1964 to 1974; the Orville G. Harrold Professorship there is named in his honor. He was elected a Fellow of the AAAS in 1954, was a Guggenheim Fellow in 1957–58, was a member of the Institute for Advanced Studies twice (1958 and 1964), and served as an Associate Secretary of the AMS from 1965 to 1976. He was married to Gladys E. Buell on June 30, 1934, and had a son, Jeffrey Buell Harrold.

At the other extreme, *Franklin Alfred Butter, Jr.* (February 1, 1910–November 27, 1972), who like Harrold also received all three of his degrees from Stanford (AB 1930, AM 1931, Ph.D. 1936) and graduated the same year as Harrold (the title of his dissertation “A Contribution to the Theory of the Arithmetic–Geometric Mean”), is more elusive. We know the basic details of his career up to 1946 thanks to his entry in the Stanford University *Annual Register* for 1945–46, when he temporarily returned to Stanford as an Acting Assistant Professor.⁴¹ After he received his Ph.D. at Stanford, he spent the year 1936–1937 at Washington University in St. Louis as a Rockefeller Science Fund Research Fellow, collaborating with Gábor Szegő in the preparation of Szegő’s classic book *Orthogonal Polynomials* (Szegő, 1939, p. vii), and then went on to teach at the University of Southern California for the next six years. He appears to have published nothing during his time at USC, however, leaving there in 1943 and spending the next three years in a series of one-year positions at the University of Wisconsin, Lawrence College and Stanford. But then his life took an unexpected turn: in 1942–43, he had been a mathematical consultant in the Engineering Department at the Douglas Aircraft Company, and this appears to have generated an interest in aerospace science (see Butter, 1945, a review of *Aircraft Analytic Geometry*). In any case, he left academia in 1946 to work in the aerospace industry: first as a mathematician and later Research Physicist at the Hughes Aircraft Company, Culver City, California, 1946–57; then an engineering specialist at Northrop Aircraft, Inc., Hawthorne, California, 1957–1961; and finally a Staff Engineer at the

Aerospace Corporation, El Segundo, California, 1961–65 (all three of these located in Los Angeles County). But he returned to teaching in 1965, when he was appointed an Associate Professor at California State College at Long Beach, and promoted to Professor there in 1970.⁴²

Finally there is *Carl Douglas Olds* (May 11, 1912–November 11, 1979), Uspensky’s last student, who received his Ph.D. in 1943, his dissertation, like Bacon’s, on number theory (“On the Number of Representations of the Square of an Integer as the Sum of an Odd Number of Squares”). He was an acting instructor at Stanford from 1935 to 1940, and an assistant professor at Purdue from 1940 to 1945, before joining the faculty of San Jose State in 1945 (perhaps as a replacement for Heaslet), where he remained for the rest of his career. He was an active member of the Mathematical Association of America, winning the 1973 Chauvenet prize for his article on the continued fraction expansion of e (Olds, 1970), and may be familiar to some because of his lovely book *Continued Fractions* (Olds, 1963).

APPENDIX B: BIBLIOGRAPHY OF J. V. USPENSKY

AJM: *American Journal of Mathematics*

AMM: *The American Mathematical Monthly*

BAMS: *Bulletin of the American Mathematical Society*

TAMS: *Transactions of the American Mathematical Society*

B.1 Thesis

1910 Some applications of continuous parameters to number theory. <https://catalog.hathitrust.org/Record/006087199> (accessed March 23, 2020).

Uspensky’s thesis studies minimization problems for pairs of linear forms $\alpha x + \beta y$ and $\gamma x + \delta y$, continuing earlier research of Voronoi and Minkowski. The main result of the thesis is an algorithm for the reduction of forms of the type $\alpha x^2 + 2bxy + cy^2$ that depend on i and $e^{2\pi i/3}$. [See Markov, Steklov and Krylov (1921, p. 4).]

B.2 Books

1937 *Introduction to Mathematical Probability*. New York: McGraw-Hill.

Reviews: F. N. David, *Biometrika* 30, 194–195; J. A. Greenwood, *The American Mathematical Monthly* 45, 471; H. T. H. Piaggio, *The Mathematical Gazette* 22, 202–204; Anonymous, *Nature* 141, 769.

1939 *Elementary Number Theory* (co-author Maxwell A. Heaslet). New York: McGraw-Hill.

⁴¹Stanford University Fifty-fifth Annual Register 1945–46. Stanford University, Published by the University. An invaluable historical resource, the registers are available online at <https://exhibits.stanford.edu/stanford-pubs/browse/annual-register-1891-1947> (last accessed January 30, 2022).

⁴²The details of Butter’s life from 1946 on have been teased from the *News and Notices* pages of *The American Mathematical Monthly*, 53 (1946), page 603; 65 (1958), page 63; 73 (1966), page 108; 78 (1971), page 107; and an announcement in the *Notices of the AMS*, 8 (1961), p. 486.

Reviews: R. Oldenburger, *Bull. Amer. Math. Soc.* 46, 202–205; G. F. Cramer, *National Mathematics Magazine* 14, 494; L. J. Mordell, *The Mathematical Gazette* 24, 295–298; H. Davenport, *Nature* 146, 418–419. MR0000236 (Reviewer: D. H. Lehmer).

1948 *Theory of Equations*. New York: McGraw-Hill.

Reviews: Garrett Birkhoff, *Science* 109, 429; T. A. Brown, *The Mathematical Gazette* 34, 140–142; L. E. Bush, *The American Mathematical Monthly* 56, 348–350; Kenneth May, *Popular Astronomy* 57, 46–47.

B.3 Papers

Unless otherwise noted, all papers up to 1925 are in Russian and all papers from 1927 onward are in English.

1906

1. On whole numbers formed with the 5th root of unity. *Matematicheskii Sbornik* 26, 1–17. [Proves $Z[\zeta_5]$ is Euclidean.]

1909

1. Note sur les nombres entiers dépendant d'une racine cinquième de l'unité. *Math. Ann.* 66, 109–112 (in French). [A reworking of his 1906 paper for the mathematical world outside of Russia.]

1912

1. Sur une série asymptotique d'Euler. *Archiv der Mathematik und Physik* 19, 370–1 (in French).

1913

1. Arithmetical proof of the Kronecker relations between the class numbers of binary quadratic forms. *Matematicheskii Sbornik* 29, 26–52. [Provides an elementary proof of the Kronecker relations. This line of investigation was later continued by Uspensky in his seven papers in the years 1925–1926.]

2. On arithmetical theorems given by Stieltjes. *Bulletin of the Mathematical Society in Kharkov* 14, 7–30.

3. On the representation of numbers by sums of squares. *Bulletin of the Mathematical Society in Kharkov* 14, 31–64.

4. On certain arithmetic theorems, I. *Bulletin of the Mathematical Society in Kharkov* 14, 88–96.

1914

1. On certain arithmetic theorems, II. *Bulletin of the Mathematical Society in Kharkov* 14, 97–99.

2. On the possibility of representing prime numbers by some of the simplest quadratic forms. Kazan. [Cited in Markov, Steklov and Krylov (1921, p. 5); no further bibliographic details are given.]

3. A rule for determining the sign in the equality $1 \cdot 2 \cdot \dots \frac{p-1}{2} \equiv \pm 1 \pmod{p}$ for a prime p of the form $4\mu + 3$. Kazan. [Cited in Markov, Steklov and Krylov (1921, p. 6); no further bibliographic details are given.]

1915

1. On the class numbers of positive Hermite forms. *Bulletin of the Imperial Academy of Sciences of St.-Petersburg* 9, 1769–1800.

1916

1. On the representation of numbers by the quadratic forms with 4 and 6 variables, I. *Bulletin of the Mathematical Society in Kharkov* 5, 81–112.

2. On the representation of numbers by the quadratic forms with 4 and 6 variables, II. *Bulletin of the Mathematical Society in Kharkov* 5, 113–147.

3. On the convergence of quadrature formulas between infinite limits. *Bulletin of the Imperial Academy of Sciences of St.-Petersburg* 10, 851–866. [Extended in Uspensky, 1928/2]

4. On the development of functions in series arising from the polynomials $e^x \frac{d^n x^n e^{-x}}{dx^n}$. *Bulletin of the Imperial Academy of Sciences of St.-Petersburg* 10, 1173–1202.

1920

1. Asymptotic formulae for numerical functions, which occur in the theory of the partition of numbers into summands. *Bulletin of the Academy of Sciences of Russia* 14, 199–218.

1921

1. About approximate expressions for the coefficients of distant terms in the development of the equation of the center into a series by the sine of multiples of the mean anomaly. *Bulletin of the Academy of Sciences of Russia* 15, 333–342.

1923

1. Note on the scientific work of A. A. Markov. *Bulletin of the Academy of Sciences of Russia* 17, 19–34.

1924

1. On a problem of Jean Bernoulli. *Bulletin of the Academy of Sciences of Russia* 18, 67–84.

1925

1. Sur les valeurs asymptotiques des coefficients de Cotes (in French). *BAMS* 31, 145–156. [Describes the asymptotic behavior of the coefficients in the Newton–Cotes method. “Attention should be called to the fact that many formulas on p. 147 ... are marred by typographical errors” (Uspensky, 1935/1, p. 382).]

2. Note sur le nombre des représentations des nombres par une somme d'un nombre pair de carrés. *Bulletin of the Academy of Sciences of the USSR* 19, 647–642.

3. Sur les relations entre les nombres des classes des formes quadratiques binaires et positives. Premier Mémoire, I. *Bulletin of the Academy of Sciences of the USSR* 19, 599–620.

4. Sur les relations entre les nombres des classes des formes quadratiques binaires et positives. Premier Mémoire, II. *Bulletin of the Academy of Sciences of the USSR* 19, 763–784.

1926

1. Sur les relations entre les nombres des classes des formes quadratiques binaires et positives. Deuxième Mémoire, I. *Bulletin of the Academy of Sciences of the USSR* 20, 25–38.

2. Sur les relations entre les nombres des classes des formes quadratiques binaires et positives. Deuxième Mémoire, II. *Bulletin of the Academy of Sciences of the USSR* 20, 175–196.

3. Sur les relations entre les nombres des classes des formes quadratiques binaires et positives. Troisième Mémoire. *Bulletin of the Academy of Sciences of the USSR* 20, 327–348.

4. Sur les relations entre les nombres des classes des formes quadratiques binaires et positives. Quatrième Mémoire. *Bulletin of the Academy of Sciences of the USSR* 20, 547–566.

5. Sur les relations entre les nombres des classes des formes quadratiques binaires et positives. Cinquième Mémoire. *Bulletin of the Academy of Sciences of the USSR*, 20 619–642.

The last seven papers rigorously derive class-number relations using elementary methods; these relations were already known but had been obtained previously using the more complex machinery of elliptic functions. For later discussion, see [Dwyer \(1937\)](#).

6. Vladimir Andreevich Steklov. *Bulletin of the Academy of Sciences of the USSR* 20, 10–11, 837–856.

1927

1. Note on the computation of roots. *AMM* 34, 130–134.

2. A curious case of the use of mathematical induction in geometry. *AMM* 34, 247–250.)

3. On a problem arising out of the theory of a certain game. *AMM* 34, 516–521.

4. On the development of arbitrary functions in series of Hermite's and Laguerre's polynomials. *Annals of Mathematics* 28, 593–619. [Uspensky says in the introduction that his paper 1916/4 “was written in Russian and published during the time when all the international relations were broken, and this may sufficiently account for the fact that it could pass unnoticed. However, as the method used by the author can successfully be applied whenever asymptotic expressions of a certain type exist, it seems worthwhile to reproduce the essential parts of this paper in a modified and simplified form.” See [Sansone \(1950\)](#).]

1928

1. On Jacobi's arithmetical theorems concerning the simultaneous representation of numbers by two different quadratic forms. *TAMS* 30, 385–404. [“Uspensky developed the elementary methods, which seem to have been used by Liouville in a series of papers published in Russian; references will be found in [this Transactions paper]. He carries his analysis up to $2s = 12$, and states that his methods enable him to prove Boulyguine's general formulae ([Hardy and Wright, 1960](#), p. 318).]

2. On the convergence of quadrature formulas related to an infinite interval. *TAMS* 30, 542–559. [L_2 convergence of Lagrange interpolation on the entire real line. [Nevai \(1986](#), p. 118) notes Uspensky's priority in dealing with the problem more than three decades earlier than other literature.]

3. On Gierster's classnumber relations. *AJM* 50, 93–122. [Applies results in his 1925–26 series of memoirs.]

4. On some new class-number relations. *Proceedings of the International Mathematical Congress* (J. C. Fields, ed.), Vol. 1, 315–317. [Written with his student, Boris Venkov]

1929

1. On the number of representations of integers by certain ternary quadratic forms. *AJM* 51, 51–60. [See [Kaplansky \(2013\)](#) for later context.]

1930

1. On the reduction of the indefinite binary quadratic forms. *BAMS* 36, 710–718.

2. On incomplete numerical functions. *BAMS* 36(10), 743–754. [Applies results in his 1925–26 series of memoirs.]

1931

1. A method for finding units in cubic orders of a negative discriminant. *TAMS* 33, 1–22. [An extension of the results in Zolotarev's classic paper of 1869.]

2. On Ch. Jordan's series for probability. *Annals of Mathematics* 32(2), 306–312. [Uspensky's paper “has been overlooked by most authors dealing with [Poisson] approximation problems” ([Deheuvels and Pfeiffer, 1988](#), pp. 671–672). Deheuvels and Pfeiffer go on to establish a relationship between Uspensky's approximation and the modern Poisson convolution semigroup approach, and exploit this to simplify and extend prior results in the literature.]

1932

1. On the problem of runs. *AMM* 39, 322–326.

1933

1. A minimum problem. *BAMS* 40, 5–10.

1934

1. On an expansion of the remainder in the Gaussian quadrature formula. *BAMS* 49, 871–876. [“In this article I shall prove that the remainder in the Gaussian formula can be expanded into a series possessing all the properties of the classical Euler–Maclaurin expansion.”]

1935

1. On the expansion of the remainder in the Newton–Cotes formula. *TAMS* 37, 381–396. [Extends work in 1934/1]

1944

1. Elementary derivation of the series for $\sin x$ and $\cos x$. (Spanish) *Math. Notae* 4, 1–10.
2. A new proof of Jacobi’s theorem. (Spanish) *Math. Notae* 4, 80–89.

1945

1. On the problem of the ruin of gamblers. (Spanish) *Publ. Inst. Mat. Univ. Nac. Litoral* 7, 155–186.
2. Sur la méthode de Laplace dans la théorie de l’attraction des ellipsoïdes homogènes. (French) *Publ. Inst. Mat. Univ. Nac. Litoral* 5, 63–71.
3. On the arithmetico–geometric means of Gauss, I–III. (Spanish) *Math. Notae* 5, 1–28, 57–88, 129–161.

1946

1. On a problem of John Bernoulli, I–IV. (Spanish) *Rev. Un. Mat. Argentina* 11, 141–154, 165–183, 239–255; 12, 10–19.

B.4 Book Reviews and Letters

1932 Review of *Studies in the Theory of Numbers*, by L. E. Dickson. *BAMS* 38, 463–465.

1940 Review of *Development of the Minkowski Geometry of Numbers* by Harris Hancock. *National Mathematics Magazine* 14, 423–424.

1944 Remarks on the History of Science in Russia. *Science* 100, 193–194 (with S. P. Timoshenko).

1946 Book Review: P. L. Chebyshev, *Collected works*. Vol. 1. *Theory of numbers*, *BAMS* 52, 50.

B.5 Problems and Solutions

Uspensky was a frequent contributor to the Problem Section of the *American Mathematical Monthly*, submitting at least eleven problems and three solutions to it: Number 3251, 34:4 (Apr. 1927), p. 216 [solution given by Uspensky 37:6 (Jun.–Jul. 1930), pp. 318–319]; Number 3278, 34:7 (Aug.–Sep. 1927), p. 381; Number 3290, 34:9 (Nov. 1927), p. 491 [reproposed by C. D. Olds – Uspensky’s student – as Number 4400, 57:6 (June 1950), p. 420]; Number 3304, 35:1 (Jan. 1928), p. 41; Number 3312, 35:3 (Mar. 1928), p. 154; Number 3343, 35:8 (Oct.

1928), p. 446 [solved by H. Langman, 36:8 (Oct. 1929), p. 450]; Number 3354, 35:10 (Dec. 1928), p. 563; Number 3389, 36:8 (Oct. 1929), p. 448; Number 3408, 37:1 (Jan. 1930), p. 38 [solution given by Uspensky 39:3 (Mar. 1932), pp. 176–177]; Number 3534, 39:2 (Feb. 1932), p. 116 [but see 68:8 (Oct. 1961), p. 814 indicating an error in the formula proposed by Uspensky, and 69:2 (Feb. 1962), pp. 172–173, showing that this was due to a simple typographical error]; solution to Number 3588 in 40:10 (Dec. 1933), p. 614 [note “Solved also by L. Zander” – his wife]

B.6 Other Writings

In addition to the papers listed above, the [Markov, Steklov and Krylov](#) report to the Russian Academy of Sciences lists two papers whose publication status was unknown to them:

- a. Application of the Poisson summation method to some types of Chebyshev polynomials.

[[Markov, Steklov and Krylov](#) (1921, p. 6) explain: “This paper should have been published in the jubilee edition of the Notices (*Izvestiya*) of the Kazan Mathematical Society. It is unknown what happened to it.” It may be this is the paper Uspensky (1927/5, p. 593) is referring to when, commenting on [Hille](#) (1926), he wrote: “From the same point of view, but in a slightly different way, the subject was treated in a paper by the author presented to the Math. Society in Kazan in 1916. Owing, however, to the political disturbances of that time, this paper never was published and was finally lost.”]

- b. On the approximate expression of the deleted terms in the expansion of the equation center in a row of sines multiple of the average anomaly. [Markov, Steklov and Krylov](#) say: “This paper should have been published in the “Notices” (*Izvestiya*) of the Perm State University, but it is unknown whether it has been printed by now.”

At the end of their report [Markov, Steklov and Krylov](#) add “Besides the aforementioned works there are preprints for the following,” and then list without date what appear to be five pedagogical books or pamphlets written by Uspensky:

1. Lectures on non-Euclidean geometry. 126 p.
2. Theory of biquadratic and cubic residues. 160 p.
3. Continuous fractions and their applications. 160 p.
4. Introduction to additive number theory. 80 p.
5. Liouville methods in number theory. 250 p.

B.7 Translations

J. Bernoulli, *Ars Conjectandi, Part 4*, translated from the Latin by Ya. V. Uspenskii, and with an introduction by A. A. Markov. St.-Petersburg: Imp. Akademiya Nauk, 1913. Reprinted with additional commentary by Yu. V. Prokhorov, 1986.

B.8 Books in Russian

Ocherk istorii logarifmov (Essay on the History of Logarithms). Petrograd: Nauchnoe knigoizdatel'stvo, 1923.

ACKNOWLEDGMENTS

We thank Alexei Borodin and Ilya Khayutin for help with the Russian literature, Stanford Librarian Ashley Jester for help with the Stanford Archives, Sunny Scott and her Stanford history experts for information about Uspensky during his time at Stanford and Stephen DeSalvo, Stewart Ethier, Jim Pitman, Eugenio Regazzini, Reinhard Siegmund-Schultze, Steve Stigler and two anonymous referees for helpful comments.

FUNDING

Persi Diaconis's research was partially supported by NSF Grant DMS-1954042.

REFERENCES

- ALEXANDERSON, G. L. (2011). Harold M. Bacon. In *Fascinating Mathematical People: Interviews and Memoirs* (D. J. Albers and G. L. Alexanderson, eds.) Chapter 3. Princeton Univ. Press, Princeton, NJ.
- BAHADUR, R. R. (1960). Some approximations to the binomial distribution function. *Ann. Math. Stat.* **31** 43–54. [MR0120675 https://doi.org/10.1214/aoms/1177705986](https://doi.org/10.1214/aoms/1177705986)
- BAHADUR, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Stat.* **37** 577–580. [MR0189095 https://doi.org/10.1214/aoms/1177699450](https://doi.org/10.1214/aoms/1177699450)
- BEATTY, S. (1926). Problem 3173. *Amer. Math. Monthly* **33** 159.
- BEATTY, S., OSTROWSKI, A., HYSLOP, J. and AITKEN, A. C. (1927). Problems and Solutions: Solutions: 3177. *Amer. Math. Monthly* **34** 159–160. [MR1521129 https://doi.org/10.2307/2298716](https://doi.org/10.2307/2298716)
- BELL, E. T. (1951). Hans Frederick Blichfeldt 1873–1945. In *Biographical Memoirs* **26** 181–189. National Academy of Science, Washington, DC.
- BENNETT, G. (1962). Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57** 33–45.
- BERNOULLI, J. (1913). *Ars Conjectandi*, Part 4. Imp. Akaemiya Nauk, St. Petersburg. Translated from the Latin by Ya. V. Uspenskii, with an introduction by A. A. Markov. Reprinted 1986, with additional notes by Yu. V. Prokhorov.
- BERNSTEIN, S. N. (1924). On a variant of Chebyshev's inequality and on the accuracy of Laplace's formula (in Russian). *Uč. Zap. Nauchn.-Issled. Kafedr Ukrainy, Otd. Mat.* **1** 38–48.
- BERTRAND, J. (1888). *Calcul des Probabilités*. Gauthier-Villars, Paris. Reprinted 1889, 1907. Reprinted, 1972 by Chelsea, New York.
- BHATIA, R. (2008). A conversation with S. R. S. Varadhan. *Math. Intelligencer* **30** 24–42. [MR2410123 https://doi.org/10.1007/BF02985734](https://doi.org/10.1007/BF02985734)
- BLACKWELL, D. (1954). On optimal systems. *Ann. Math. Stat.* **25** 394–397. [MR0061776 https://doi.org/10.1214/aoms/1177728798](https://doi.org/10.1214/aoms/1177728798)
- BLICHFELDT, H. F. (1900). On a certain class of groups of transformation in space of three dimensions. *Amer. J. Math.* **22** 113–120. [MR1505825 https://doi.org/10.2307/2369749](https://doi.org/10.2307/2369749)
- BOHR, H. and MOLLERUP, J. (1922). *Laerebog i Matematisk Analyse* **3**. In Danish.
- BORWEIN, J. M., CHOI, K.-K. S. and PIGULLA, W. (2005). Continued fractions of tails of hypergeometric series. *Amer. Math. Monthly* **112** 493–501. [MR2142600 https://doi.org/10.2307/30037519](https://doi.org/10.2307/30037519)
- BURNSIDE, W. (1928). *Theory of Probability*. Cambridge Univ. Press, Cambridge. 2nd ed., 1936.
- BUTTER, F. A. JR. (1945). Recent Publications: Aircraft Analytic Geometry. *Amer. Math. Monthly* **52** 338–340. [MR1526233 https://doi.org/10.2307/2305297](https://doi.org/10.2307/2305297)
- CASSELS, J. W. S. and VAUGHAN, R. C. (1985). Obituary: Ivan Matveevich Vinogradov. *Bull. Lond. Math. Soc.* **17** 584–600. [MR0813744 https://doi.org/10.1112/blms/17.6.584](https://doi.org/10.1112/blms/17.6.584)
- CHAMBERLAIN, L. (2006). *Lenin's Private War*. St. Martin's Press, New York.
- COCHRAN, W. G. (1952). The χ^2 test of goodness of fit. *Ann. Math. Stat.* **23** 315–345. [MR0049531 https://doi.org/10.1214/aoms/1177729380](https://doi.org/10.1214/aoms/1177729380)
- COOLIDGE, J. L. (1925). *An Introduction to Mathematical Probability*. Clarendon Press, Oxford.
- CRAIG, C. C. (1933). On the Tchebychev inequality of Bernstein. *Ann. Math. Stat.* **4** 94–102.
- CRAMÉR, H. (1976). Half a century with probability theory: Some personal recollections. *Ann. Probab.* **4** 509–546. [MR0402837 https://doi.org/10.1214/aop/1176996025](https://doi.org/10.1214/aop/1176996025)
- DAVID, F. N. (1938). Review of introduction to mathematical probability by J. V. Uspensky. *Biometrika* **30** 194–195.
- DE MOIVRE, A. (1967). *The Doctrine of Chances: A Method of Calculating the Probabilities of Events in Play*. Chelsea Publishing Co., New York. [MR0216923](https://doi.org/10.2307/2305297)
- DEHEUVELS, P. and PFEIFER, D. (1988). On a relationship between Uspensky's theorem and Poisson approximations. *Ann. Inst. Statist. Math.* **40** 671–681. [MR0996692 https://doi.org/10.1007/BF00049425](https://doi.org/10.1007/BF00049425)
- DEMIDOV, S. S. (2015). World War I and mathematics in “the Russian world”. *Czas. Tech.* **112** 77–92.
- DESALVO, S. (2021). Will the real Hardy–Ramanujan formula please stand up? *Integers* **21** A88. [MR4304355](https://doi.org/10.2307/2305297)
- DEWITT, N. (1961). *Education and Professional Employment in the U.S.S.R.* National Science Foundation, Washington, DC.
- DIACONIS, P. and GRAHAM, R. (2012). *Magical Mathematics: The Mathematical Ideas That Animate Great Magic Tricks*. Princeton Univ. Press, Princeton, NJ. [MR2858033](https://doi.org/10.2307/2305297)
- DIACONIS, P., GRAHAM, R. L. and KANTOR, W. M. (1983). The mathematics of perfect shuffles. *Adv. in Appl. Math.* **4** 175–196. [MR0700845 https://doi.org/10.1016/0196-8858\(83\)90009-X](https://doi.org/10.1016/0196-8858(83)90009-X)
- DIACONIS, P. and ZABELL, S. (1991). Closed form summation for classical distributions: Variations on a theme of de Moivre. *Statist. Sci.* **6** 284–302. [MR1144242](https://doi.org/10.2307/2305297)
- DICKSON, L. E. (1947). Hans Frederik Blichfeldt: 1873–1945. *Bull. Amer. Math. Soc.* **53** 882–883.
- DUDLEY, R. M. (1987). Some inequalities for continued fractions. *Math. Comp.* **49** 585–593. [MR0906191 https://doi.org/10.2307/2008331](https://doi.org/10.2307/2008331)
- DUTKA, J. (1981). The incomplete beta function—a historical profile. *Arch. Hist. Exact Sci.* **24** 11–29. [MR0618150 https://doi.org/10.1007/BF00327713](https://doi.org/10.1007/BF00327713)
- DWYER, W. A. (1937). On Certain Fundamental Identities Due to Uspensky. *Amer. J. Math.* **59** 290–294. [MR1507239 https://doi.org/10.2307/2371411](https://doi.org/10.2307/2371411)
- ERMOLAEVA, N. (1997). Uspensky, Yakov Viktorovich. In *Russian Diaspora: The Golden Book of Emigration, the First Third of the 20th Century, an Encyclopedic Biographical Dictionary* (V. V. Shelokhaeva and N. I. Kanishcheva, eds.) ROSSPEN, Moscow.

- ETHIER, S. N. (2010). *The Doctrine of Chances. Probability and Its Applications* (New York). Springer, Berlin. MR2656351 <https://doi.org/10.1007/978-3-540-78783-9>
- ETHIER, S. N. and KHOSHNEVISAN, D. (2002). Bounds on gambler's ruin probabilities in terms of moments. *Methodol. Comput. Appl. Probab.* **4** 55–68. MR1918933 <https://doi.org/10.1023/A:1015705430513>
- FELLER, W. (1950). *An Introduction to Probability Theory and Its Applications*. Wiley, New York. 2nd ed. 1957, 3rd ed. 1968.
- FINKEL, S. (2003). Purging the public intellectual: The 1922 expulsions from Soviet Russia. *Russ. Rev.* **62** 589–613.
- FISHER, A. (1922). *The Mathematical Theory of Probabilities & Its Application to Frequency Curves & Statistical Methods*. Macmillan, New York. 1st ed., 1915.
- FISHER, R. A. (1925). *Statistical Methods for Research Workers*. Oliver & Boyd, Edinburgh. 14th (posthumous) ed., 1970, Hafner, New York.
- FORSYTH, C. H. (1925). Simple derivations of the formulas for the dispersion of a statistical series. *Amer. Math. Monthly* **31** 190–196.
- FRY, T. C. (1965). *Probability and Its Engineering Uses*, 2nd ed. D. Van Nostrand Co., Princeton, NJ–Toronto–London. MR0181027
- GILLILAND, D., LEVENTAL, S. and XIAO, Y. (2007). A note on absorption probabilities in one-dimensional random walk via complex-valued martingales. *Statist. Probab. Lett.* **77** 1098–1105. MR2395066 <https://doi.org/10.1016/j.spl.2007.01.016>
- GRAHAM, R. L. (1963). On a theorem of Uspensky. *Amer. Math. Monthly* **70** 407–409. MR0148555 <https://doi.org/10.2307/2311859>
- GRAHAM, R. L., LIN, S. and LIN, C. S. (1978). Spectra of numbers. *Math. Mag.* **51** 174–176. MR0491580 <https://doi.org/10.2307/2689998>
- GRAHAM, R. and O'BRYANT, K. (2005). A discrete Fourier kernel and Fraenkel's tiling conjecture. *Acta Arith.* **118** 283–304. MR2168767 <https://doi.org/10.4064/aa118-3-4>
- HALD, A. (1990). *A History of Probability and Statistics and Their Applications Before 1750*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Wiley, New York. MR1029276 <https://doi.org/10.1002/0471725161>
- HALD, A. (1998). *A History of Mathematical Statistics from 1750 to 1930*. Wiley Series in Probability and Statistics: Texts and References Section. Wiley, New York. MR1619032
- HARDY, G. H. and RAMANUJAN, S. (1918). Asymptotic formulae in combinatory analysis. *Proc. Lond. Math. Soc.* **17** 75–115. MR1575586 <https://doi.org/10.1112/plms/s2-17.1.75>
- HARDY, G. H. and WRIGHT, E. M. (1960). *The Theory of Numbers*, 4th ed. Clarendon Press, Oxford.
- HARTMAN, E. P. (1970). *Adventures in Research: A History of Ames Research Center 1940–1965*. National Aeronautics and Space Administration, Washington, DC.
- HEYDE, C. C. and SENETA, E. (1977). *I. J. Bienaymé. Statistical Theory Anticipated. Studies in the History of Mathematics and Physical Sciences*, No. 3. Springer, New York. MR0462888
- HILLE, E. (1926). On Laguerre's series. *Proc. Natl. Acad. Sci. USA* **12** 261–265, 265–269, 348–352.
- JACKSON, D. (1935). Mathematical principles in the theory of small samples. *Amer. Math. Monthly* **72** 344–364.
- JACKSON, B. W. (no date). A History of the Mathematics Department at San Jose State. (accessed January 30, 2022).
- JELLISON, C. (1997). The prisoner and the professor. *Stanford Mag.* **72**. March/April issue, <https://stanfordmag.org/contents/the-prisoner-and-the-professor>.
- KAPLANSKY, I. (2013). Integers uniquely represented by certain ternary forms. In *The Mathematics of Paul Erdős I* (R. L. Graham et al., eds.) 71–79. Springer, New York.
- KARATSUBA, A. A. (1981). Ivan Matveevich Vinogradov (on his ninetieth birthday). *Russian Math. Surveys* **36** 1–17.
- KRANDICK, W. and MEHLHORN, K. (2006). New bounds for the Descartes method. *J. Symbolic Comput.* **41** 49–66. MR2194885 <https://doi.org/10.1016/j.jsc.2005.02.004>
- LORENTZ, G. G. (2002). Mathematics and politics in the Soviet Union from 1928 to 1953. *J. Approx. Theory* **116** 169–223. MR1911079 <https://doi.org/10.1006/jath.2002.3670>
- MALYSHEV, A. V. and FADDEEV, D. K. (1961). Boris Alekseevich Venkov (on his sixtieth birthday). *Uspekhi Mat. Nauk* **16** 235–240.
- MARKOV, A. A. (1888). *Table des valeurs de l'intégrale $\int_x^\infty e^{-t^2} dt$* . Imperial Academy of Sciences, St. Petersburg.
- MARKOV, A. A. (1900). *Ischislenie Veroiatnostei (Calculus of Probabilities)*. Tipografija Imperatorskoj Akademii Nauk, Sanktpeterburg. Three subsequent editions: 2nd ed., 1908, and 3rd ed., 1913; 4th (posthumous) ed., 1924, Gosudarstvennoe Izdatel'stvo, Moscow. German translation of 2nd ed.: *Wahrscheinlichkeitsrechnung* (Heinrich Liebmann, trans.), Teubner, Leipzig and Berlin, 1912.
- MARKOV, A., STEKLOV, V. and KRYLOV, N. (1921). Note on the scientific work of Petrograd University Professor Yakov Viktorovich Uspensky. *Izv. Rossiyskoy Akad. Nauk (Ser. 6)* **15** 4–6.
- MCKAY, B. D. (1989). On Littlewood's estimate for the binomial distribution. *Adv. in Appl. Probab.* **21** 475–478. MR0997736 <https://doi.org/10.2307/1427172>
- MILLER, G. H. (1970). Blichfeldt, Hans Frederick. In *Dictionary of Scientific Biography* (C. C. Gillispie, ed.) **2**. Charles Scribner's Son, New York. Reprinted 2008 in the ebook *Complete Dictionary of Scientific Biography*.
- MOLINA, E. C. (1944). Arne Fisher, 1887–1944. *J. Amer. Statist. Assoc.* **39** 251.
- MUNROE, M. E. (1951). *Theory of Probability*. McGraw-Hill, Inc., New York–Toronto–London. MR0040604
- NAZAROV, A. I. and SINKEVICH, G. I. (2019). History of Leningrad mathematics in the first half of the 20th century. *ICCM Not.* **7** 47–63. MR3995130 <https://doi.org/10.4310/ICCM.2019.v7.n2.a5>
- NEVAI, P. (1986). Géza Freud, orthogonal polynomials and Christoffel functions. A case study. *J. Approx. Theory* **48** 3–167. MR0862231 [https://doi.org/10.1016/0021-9045\(86\)90016-X](https://doi.org/10.1016/0021-9045(86)90016-X)
- OLDS, C. D. (1963). *Continued Fractions*. Random House, New York. MR0146146
- OLDS, C. D. (1970). The simple continued fraction expansion of e . *Amer. Math. Monthly* **77** 968–974. MR1536091 <https://doi.org/10.2307/2318113>
- ONDAR, K. O. (1981). *The Correspondence Between A. A. Markov and A. A. Chuprov on the Theory of Probability and Mathematical Statistics*. Springer, New York. Translated by Charles and Margaret Stein.
- OSTROWSKI, A. M. (1950). Note on Vincent's theorem. *Ann. of Math.* (2) **52** 702–707. MR0038481 <https://doi.org/10.2307/1969443>
- PEARSON, K. (1931). Historical note on the distribution of the standard deviations of samples of any size drawn from an indefinitely large normal parent population. *Biometrika* **23** 416–418.
- PEARSON, E. (1990). *'Student': A Statistical Biography of William Sealy Gosset*. Clarendon Press, Oxford. R. L. Plackett and G. A. Barnard, eds.
- PEIZER, D. B. and PRATT, J. W. (1968). A normal approximation for binomial, F , beta, and other common, related tail probabilities. *I. J. Amer. Statist. Assoc.* **63** 1416–1456. MR0235650
- PERLMAN, M. D. and WICHURA, M. J. (1975). Sharpening Buffon's needle. *Amer. Statist.* **29** 157–163. MR0464479 <https://doi.org/10.2307/2683484>
- PÓLYA, G., SZEGŐ, G. and YOUNG, D. H. (1947). Memorial Resolution: James Victor Uspensky. Stanford University Archives.

- RASHDALL, H. (1895). *The Universities of Europe in the Middle Ages. Volume 1: Salerno—Bologna—Paris*. Clarendon Press, Oxford.
- REEDS, J. A., DIFFIE, W. and FIELD, J. V. (2015). *Breaking Teleprinter Ciphers at Bletchley Park*. Wiley–IEEE Press, Piscataway, NJ.
- RIETZ, H. L. (1924). *Handbook of Mathematical Statistics*. Houghton Mifflin, Boston, MA.
- ROUCHÉ, E. (1888). Sur un problème relatif à la durée du jeu. *C. R. Acad. Sci.* **116** 47–49.
- ROYDEN, H. (1989). The history of the mathematics department at Stanford. In *A Century of Mathematics in America, Part II* (P. Duren, ed.) 237–278. Amer. Math. Soc., Providence, RI.
- ROYDEN, H., LYMAN, R., PHILLIPS, R., SUNSERI, M. and WINBIGLER, H. D. (1992). Memorial resolution: Harold M. Bacon (1907–1992). Department of Special Collections and University Archives. Available at <https://purl.stanford.edu/ws360gk0273> (accessed February 29, 2020).
- SANSONE, G. (1950). La formula di approssimazione asintotica dei polinomi di Tchebychef-Laguerre col procedimento di J. V. Uspensky. *Math. Z.* **53** 97–105. [MR0037940 https://doi.org/10.1007/BF01162402](https://doi.org/10.1007/BF01162402)
- SEABASS, R. A. (2002). Harvard Lomax. In *MAM2006: Markov Anniversary Meeting* (A. N. Langville and W. J. Stewart, eds.) 169–172. The National Academies Press.
- SEAL, H. (1966). The random walk of a simple risk business. *ASTIN Bull. J. IAA* **4** 19–28.
- SENETA, E. (2006). Markov and the the creation of Markov chains. In *MAM2006: Markov Anniversary Meeting* (A. N. Langville and W. J. Stewart, eds.) 1–20. Bosen Books, Raleigh, NC.
- SENETA, E. (2013). A tricentenary history of the law of large numbers. *Bernoulli* **19** 1088–1121. [MR3102545 https://doi.org/10.3150/12-BEJSP12](https://doi.org/10.3150/12-BEJSP12)
- SHEYNIN, O. (2005). Jakob Bernoulli: On the Law of Large Numbers. Unpublished translation.
- SIEGMUND-SCHULTZE, R. (2006). Probability in 1919/20: The von Mises-Pólya controversy. *Arch. Hist. Exact Sci.* **60** 431–515. [MR2238933 https://doi.org/10.1007/s00407-006-0112-x](https://doi.org/10.1007/s00407-006-0112-x)
- STANFORD (1992). Long-time calculus professor Harold Bacon dies at 85. Stanford University News Service.
- STIGLER, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty Before 1900*. The Belknap Press of Harvard Univ. Press, Cambridge, MA. [MR0852410](https://doi.org/10.1007/BF01174358)
- STRUTT, J. W. (1894). *The Theory of Sound, Volume 1*, 2nd ed. Macmillan, London and New York. (3rd Baron Rayleigh).
- SVIDERSKAYA, G. E. (2002). Rodion Osievich Kuzmin.
- SYLLA, E. (2006). *Jacob Bernoulli: The Art of Conjecturing*. Johns Hopkins Univ. Press, Baltimore, MD. Translation of Jacob Bernoulli's *Ars conjectandi*, with an introduction and notes, by Edith Dudley Sylla.
- SZEGÖ, G. (1939). *Orthogonal Polynomials. American Mathematical Society Colloquium Publications, Vol. 23*. Amer. Math. Soc., New York. [MR0000077](https://doi.org/10.1007/BF01174358)
- TIMOSHENKO, S. (1947). Typescript. Unpublished manuscript, presumably prepared shortly after Uspensky's death and in the Stanford University Archives.
- TIMOSHENKO, S. (1968). *As I Remember: The Autobiography of Stephen P. Timoshenko*. Van Nostrand, Princeton, NJ. Translated from the Russian by Robert Addis.
- TODHUNTER, I. (1865). *A History of the Mathematical Theory of Probability*. Macmillan, Cambridge and London. Reprinted, 1949 and 1964, Chelsea, New York. ["A work of true learning, beyond criticism"—John Maynard Keynes.].
- USPENSKY, J. V. (1906). On whole numbers formed with the 5th root of unity. *Mat. Sb.* **26** 1–17.
- USPENSKY, J. V. (1920). Asymptotic formulae for numerical functions which occur in the theory of the partition of numbers into summands. *Bull. Acad. Sci. Russ.* **14** 199–218.
- USPENSKY, J. V. (1927). On a problem arising out of the theory of a certain game. *Amer. Math. Monthly* **34** 516–521. [MR1521314 https://doi.org/10.2307/2299838](https://doi.org/10.2307/2299838)
- USPENSKY, J. V. (1937). *Introduction to Mathematical Probability*. McGraw-Hill, New York.
- USPENSKY, J. V. (1948). *Theory of Equations*. McGraw-Hill, New York.
- USPENSKY, J. V. and HEASLET, M. A. (1939). *Elementary Number Theory*. McGraw-Hill, Inc., New York. [MR0000236](https://doi.org/10.1007/BF01174358)
- USPENSKY, J. V. and VENKOV, B. (1928). On some new class-number relations. *Bull. Acad. Sci. Russ.* **1** 315–317.
- VENKOV, B. A. and NATANSON, I. P. (1949). Rodion Osievich Kuzmin (1891–1949) (obituary). *Uspekhi Mat. Nauk* **4** 148–155.
- VERGER, J. (2003). Teachers. In *A History of the University in Europe. Volume 1: Universities in the Middle Ages* (H. D. Ridder-Symoens, ed.) 144–168. Cambridge Univ. Press, Cambridge.
- WALL, H. S. (1948). *Analytic Theory of Continued Fractions*. D. Van Nostrand, New York, NY. [MR0025596](https://doi.org/10.1007/BF01174358)
- WENKOV, B. A. (1931). Über die Klassenanzahl positiver binärer quadratischer Formen. *Math. Z.* **33** 350–374. [MR1545216 https://doi.org/10.1007/BF01174358](https://doi.org/10.1007/BF01174358)
- WYTHOFF, W. A. (1907). A modification of the game of Nim. *Nieuw Arch. Wiskd.* (5) **7** 199–202.
- ZABELL, S. L. (2008). On student's 1908 article "The probable error of a mean". *J. Amer. Statist. Assoc.* **103** 1–7.