

# Meanders\*

*S. K. Lando and A. K. Zvonkin*

Five miles meandering with a mazy  
motion.

*S. T. Coleridge, Kubla Khan*

By slow Meander's margent green  
And in the violet-embroidered vale.

*John Milton, Comus*

## 1. Introduction. Formulation of the problem

**1.1. Definition.** Let us fix a straight line on the plane and  $2n$  points on it. Consider a simple connected not self-intersecting closed curve intersecting the line in exactly those points. The equivalence class of such curves with respect to isotopies of the plane leaving the line fixed is called a (closed) *meander* of order  $n$ .

**Notation.** We denote the number of meanders of order  $n$  by  $M_n$  and call it the  $n$ -th meandric number. We set  $M_0 = 1$ . Figure 1.1 shows that  $M_1 = 1$ ,  $M_2 = 2$ ,  $M_3 = 8$ .

The present paper is devoted to the problem of enumerating meanders. It should be noted that little is known about the sequence  $\{M_n\}$ . Even its asymptotics is unknown; some estimates will be given below.

**1.2. Motivation.** The problem of enumerating meanders was formulated by V. I. Arnol'd in 1987 [1]. The term "meander" was also proposed by him. An equivalent problem was independently treated by A. Phillips [13]. P. Rosenstiehl [15] introduced analogous objects and corresponding "planar" (meandric) permutations.

Paper [1] examines "nonclosed" meanders. Our situation corresponds to an odd number of intersection points, and our  $n$ -th meandric number is the  $(2n - 1)$ -th meandric number in the sense of Arnol'd.

The problem of enumerating meanders is actually that of enumerating the possible mutual positions of two circles on a sphere. It is closely related to Hilbert's sixteenth problem concerning the number of ovals of an algebraic

\* Previously unpublished. Manuscript received December 1, 1990.

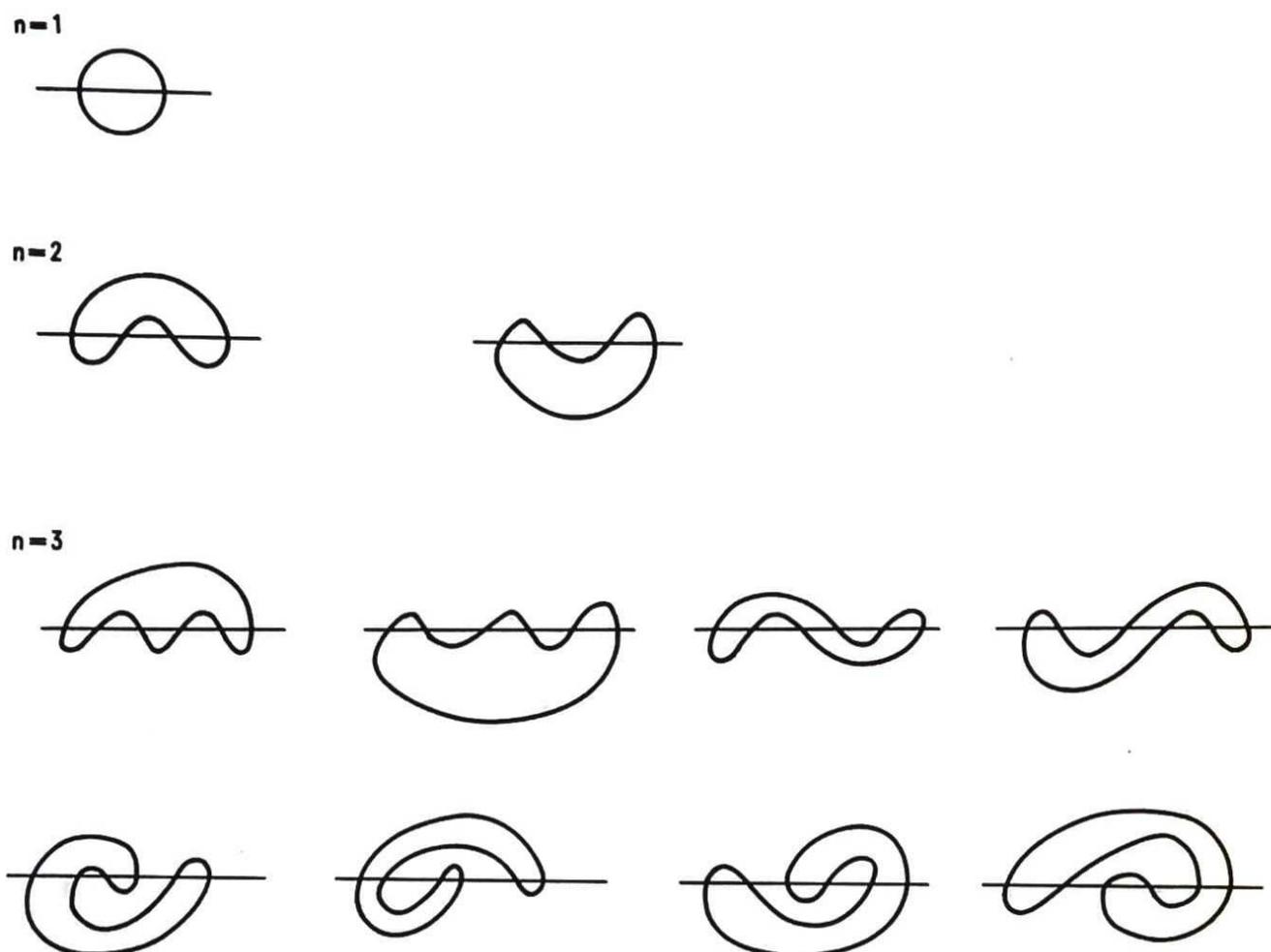


Figure 1.1. Meanders of order  $n = 1, 2, 3$ .

curve on the plane. The “projective” meanders examined in [1] (i.e., meanders on a projective plane) serve as a tool for analyzing the differential-geometrical properties of the manifolds of zeroes of hyperbolic polynomials. Arnol’d refers to these properties as “projective topology”.

The problem is also very attractive because it is unexpectedly difficult. (However, any problem may seem difficult until it is solved.) Obviously it is very beautiful. The feeling of beauty, normally subjective and hard to formalize, receives in this case a striking objective confirmation. Paper [13] enumerates a class of mazes exactly equivalent to meanders, and contains a vast catalogue of mazes of this class occurring in the architecture, painting, and decorative art of many cultures and epochs.

The problem of meanders demonstrates various links with different fields of mathematics. For instance, the best upper estimate for  $M_n$  we know of (see Section 5) has been obtained by means of a method close to the theory of

formal grammars. In Section 7 we describe the relation of the problem to enumerating Feynman diagrams in some models of quantum field theory. This section is based on a private communication by V. Kazakov.

### 1.3. Versions of the formulation of the problem

- (1) By omitting the condition that the curve be closed we obtain the well-known (and also still unsolved) problem of folding a strip of stamps (see [9], [10]). The number of intersection points may, in this case, be odd.
- (2) We get another formulation of this problem if we require that one end of the nonclosed curve come from "infinity on the left" and the second pass to "infinity on the right" (meanders in the sense of Arnol'd).
- (3) By closing the ends of the straight line into a circle we arrive at the situation in which some of the meanders become indistinguishable as they are obtained from one another by rotation. The problem again is to enumerate the different pictures of this kind.
- (4) If in the formulation of the problem we give up the requirement of connectedness and consider a system of several nonintersecting circles crossing the straight line at the same points, we come to the problem of enumerating *systems of meanders* with a given number of components.

### 1.4. Acknowledgments

V. I. Arnol'd drew our attention to the problem of meanders and supported our work. We are grateful to Professor A. Phillips for his letter [12] containing many interesting ideas and a great deal of useful information. V. Kazakov and I. Kostov kindly gave us permission to quote their unpublished result (see Section 7). A. Pinchuk and D. Ivanov helped us with the programming. We are also grateful to S. Chmutov, A. Kulakov, S. Tabachnikov, and M. Shubin for fruitful discussions.

## 2. Definitions, notation, formulation of the results

**2.1. Definition.** By omitting in the definition of a meander the requirement of connectedness of the curve under consideration, we obtain the definition of a *system of meanders*.

### 2.2. Overview of the main results

In Section 3 we describe an algorithm for enumerating all the meanders of order up to  $n$ . The complexity of this algorithm is proportional to the number of meanders and, therefore, as follows from the results of Section 4, it grows

exponentially as  $n$  increases. It would be interesting to find an algorithm for computing meandric numbers that has polynomial complexity. In the same section, the table of values of the first fourteen meandric numbers is given.

In Section 4 some simple properties of the sequence of meandric numbers are described. Proposition 4.4 states that there exists a number  $A_M$  such that the sequence of meandric numbers grows more slowly than  $(A_M)^n$  and more quickly than  $A^n$  for any  $A < A_M$ . Actually this result was communicated to Arnol'd by J. A. Reeds and L. A. Shepp in 1988.

In Section 5 the notion of an irreducible system of meanders is introduced and systems of this sort are examined. In fact the techniques of formal grammars allow us to write a generating function for the number of such systems and to calculate the convergence radius of the corresponding power series. This radius turns out to be equal to  $((4 - \pi)/\pi)^2$ , which yields an upper estimate for the number  $A_M$  defined above.

We combine these last two results in the following

**Theorem.** *There exists a number  $A_M$  such that, for any  $0 < A < A_M$ ,  $A^n < M_n < (A_M)^n$  for every large enough  $n$ , and  $A_M < (\pi/(4 - \pi))^2$ .*

According to the information contained in [13] the same estimate was derived by A. Lehman in connection with the problem of folding a strip of stamps.

In Section 6 we describe a technique for the empirical study of sequences, which is based on multipoint Padé approximations. The results of applying the technique to the investigation of the sequence of meandric numbers allow us to formulate the hypothesis that

$$M_n \sim \text{const} \cdot (A_M)^n \cdot n^{-7/2},$$

where  $A_M = 12.26 \dots$

In Section 7 we describe the relation of the problem of meanders to a model of quantum field theory and write a corresponding matrix integral. General information on matrix integrals and methods for their calculation is also given.

### 2.3. Encoding

Below we will use two ways of representing systems of meanders.

To each system of meanders there corresponds an ordered pair of regular bracket structures. Namely, the straight line divides the plane into two semi-planes: upper and lower. The part of the curve in each of these semi-planes determines a regular bracket structure: the left bracket corresponds to the left end of the arc, and the right bracket corresponds to the right end. And vice

versa, a system of meanders can be unambiguously recovered from an ordered pair of regular bracket structures.

Denoting the left bracket by  $x$  and the right one by  $y$ , we represent a regular bracket structure by a word in the alphabet  $\{x, y\}$ . The language of regular bracket structures is commonly called the Dyck language.

Now each system of meanders of order  $n$  can be presented by a word of the form  $w_1 b w_2$  in the alphabet  $\{x, y, b\}$ , where  $w_1, w_2$  are two words of length  $2n$  in the Dyck language, while the symbol  $b$  is intended for dividing the upper and lower bracket structures.

We will also make use of the encoding system proposed by J. A. Reeds (see [13]). Each system of meanders of order  $n$  can be represented by a word of  $2n$  letters in the alphabet  $\{a, b, c, d\}$ . To each of the  $2n$  points of intersection of the curve and the straight line there corresponds one of the letters  $a, b, c, d$  depending on which bracket, the right one or the left one, the point corresponds to in the upper and lower bracket structures according to the following rule:

$$\left( \begin{array}{l} \rightarrow a; \\ \left( \rightarrow b; \right) \end{array} \right) \rightarrow c; \left( \right) \rightarrow d.$$

**Example.** The system of meanders in Figure 2.1 is represented by the word  $xxxxyyyxxyybxxyxxxxyyyy$  in the language of bracket structures, and by the word  $abaccdabdd$  in the language of Reeds.

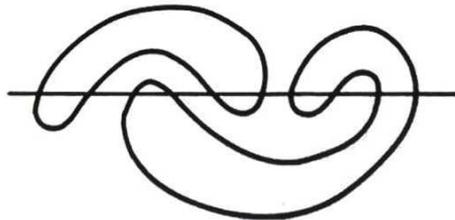


Figure 2.1.

#### 2.4. Catalan numbers

There exist many different interpretations of Catalan numbers. We call  $n$ -th *Catalan number* the number of regular bracket structures constructed of  $n$  pairs of brackets, i.e. the number of words of length  $2n$  in the Dyck language. We denote this number by  $\text{Cat}_n$ . Let  $\text{Cat}_0 = 1$ .

It is well known that

$$\text{Cat}_n = \frac{1}{n+1} \binom{2n}{n}. \quad (2.1)$$

Here is the beginning of the Catalan sequence: 1, 1, 2, 5, 8, 14, 42, . . . .  
 (2.1) implies, using the Stirling formula, that

$$\text{Cat}_n \sim \text{const} \cdot \frac{4^n}{n^{3/2}}. \tag{2.2}$$

We denote the generating function for the Catalan numbers by  $\text{Cat}(x)$ :

$$\text{Cat}(x) = 1 + x + 2x^2 + 5x^3 + \dots$$

**3. An algorithm for the enumeration.**  
**The table of meandric numbers**

We have not succeeded in finding an efficient algorithm for calculating the meandric numbers. Therefore, to obtain the meandric numbers by computer, we used the algorithm of enumerating meanders described below.

**3.1. The tree of meanders**

The enumerating algorithm is based on the representation of the whole set of meanders as an (infinite) tree where each of the meanders occurs once and only once. The vertices of the tree are meanders and the root is the unique meander of order 1. The vertices of level  $n$  correspond to the meanders of order  $n$ .

To describe the tree of meanders we will use the representation of a meander as a pair of regular bracket structures. Thus:

- (1) (The root) The meander  $xybxy$  is the only vertex of level 1.
- (2) (Pulling) If  $w^n = w_1xw_2yw_3bw$ , where  $w_1, w_2$  and  $w_3$  are Dyck words, is a vertex of level  $n$ , it is connected by an edge to the vertex  $xxw_1yw_2yw_3bxyw$  of level  $n + 1$  (see Figure 3.1). We say that  $w^{n+1}$  is obtained from  $w^n$  by pulling.



Figure 3.1.

- (3) (Restricted left shift) Besides the vertex described in (2), the vertex  $w^n$  is connected with some other vertices of level  $n + 1$ . The latter are obtained from a meander of order  $n + 1$  described in (2) by the operation of restricted shift. Namely, if the vertex  $w^n$  of level  $n$  is connected by an

edge with the vertex  $w^{n+1} = xw_1yw_2bxw_3yw_4$  of level  $n + 1$ , where  $w_1$ ,  $w_2$ ,  $w_3$  and  $w_4$  are Dyck words and  $w_3$  is not empty, then  $w^n$  is also connected by an edge with the vertex  $w_1xw_2yb_3xw_4y$  (see Figure 3.2), which is said to be obtained from  $w^{n+1}$  by left shift.

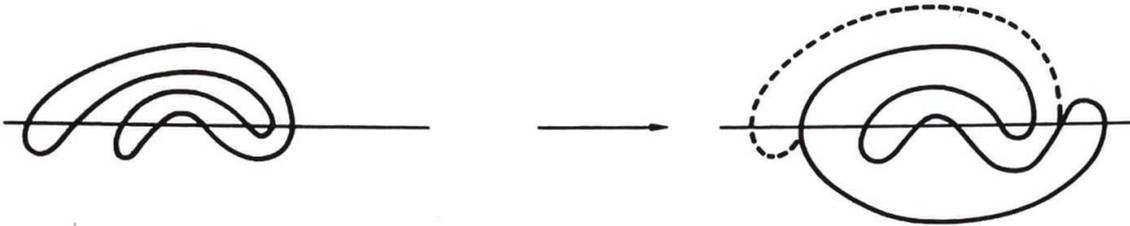


Figure 3.2.

**Proposition.** *The above description determines a tree of meanders in which each meander occurs once and only once.*

**Proof.** Let  $n > 0$ ,  $w^{n+1} = w_1bw_2$ . If the word  $w_2$  may be represented as  $w_2 = xyw'_2$ , then the word  $w_1$  may be represented as  $w_1 = xxw_1^1yw_1^2yw_1^3$ , and the vertex  $w^{n+1}$  is connected by an edge with the vertex  $w^n = w_1^1xw_1^2yw_1^3bw'_2$  as  $w^{n+1}$  is obtained from  $w^n$  by pulling. If  $w_2$  cannot be represented as  $w_2 = xyw'_2$ , we shall act on it (and on the meander as a whole) by right shifts until we obtain the word  $xy\tilde{w}_2$ . The meander  $w^{n+1}$  is connected with the same vertex of level  $n$  as the one obtained as the result of this transformation.

### 3.2. The enumerating algorithm

To enumerate all the meanders of order  $n$ , consider a subtree of the tree of meanders composed of all the vertices of level at most  $n$ . In order to enumerate all the vertices of level  $n$  it suffices to use the standard algorithm of traversing the tree.

**Notes.** (1) It is convenient to represent a meander of order  $n$  in a program by a pair of arrays of  $2n$  integers. The  $k$ -th number in the first (respectively, second) array is the number of the point to which the  $k$ -th point is connected by the upper (lower) arc.

(2) The given algorithm requires very little memory for the data: the size of the memory used is proportional to  $n^2$ .

(3) The efficiency of the algorithm is proportional to the number of meanders of order  $n$ . Therefore the working time of the algorithm increases exponentially with respect to  $n$  (see 4.2).

(4) Insignificant modifications of the algorithm allow us to compute the  $(n + 1)$ -th meandric number without enumerating the meanders of order  $n + 1$  but only those of order  $n$ .

**3.3. Table of meandric numbers**

The values of the first 14 meandric numbers are given below. We computed the first ten by means of the above algorithm. The remaining values were calculated by J. A. Reeds and L. A. Shepp.

$n$	0	1	2	3	4	5	6	7	8	9	10
$M_n$	1	1	2	8	42	262	1828	13820	110954	933458	8152860
$n$	11			12			13			14	
$M_n$	73424650			678390116			6405031050			61606881612	

**4. Simple properties of meandric numbers**

**4.1. Proposition.** *The sequence  $M_n$  increases monotonically,  $M_n < M_{n+1}$ ,  $n = 1, 2, \dots$*

**Proof.** The “pulling” operation (see 3.1) transforming the word  $w_1 x w_2 y w_3 b w$  into the word  $x x w_1 y w_2 y w_3 b x y w$  generates, for each meander of order  $n$ , one or more meanders of order  $n + 1$ . All meanders thus generated are different, and this proves the proposition.

**4.2. Proposition.**  $Cat_n \leq M_n \leq (Cat_n)^2$ .

**Proof.**  $(Cat_n)^2$  is the total number of systems of meanders of order  $n$ , which yields the upper bound. In order to prove the lower bound it suffices to note that each (upper) bracket structure can be completed to a meander by a suitable (lower) bracket structure. The latter statement is proved by induction, applying pulling and left shift (see 3.1).

**Corollary.**  $const \cdot 4^n/n^{3/2} \lesssim M_n \lesssim const \cdot 4^{2n}/n^3$ .

**4.3. Proposition.** *If  $n = p^k$  is a degree of a prime number  $p$  then  $M_n = M_{p^k} \equiv 2 \pmod{p}$ .*

**Notes.** (1) This fact was discovered by J. A. Reeds (see [13]).

(2) The same statement is true for Catalan numbers and the number of irreducible systems of meanders (see Section 5). Its proof remains unchanged.

**Proof.** On the set of meanders of order  $n$ , the group  $Z_{2n}$  acts by cyclic shifts to the left (see 3.1). When  $n = p^k$ , if an orbit has order  $m$  not divisible by  $p$ , then  $m$  is equal to 1 or 2. When  $n > 1$  the group has no fixed points and there exists a unique orbit of order 2 (see Figure 4.1).



Figure 4.1.

**4.4. Proposition.** *There exists a number  $A_M$  such that for any  $A$ ,  $0 < A < A_M$ ,  $A^n < M_n \leq (A_M)^n$  for all sufficiently large  $n$ .*

To prove the proposition we will need the following

**Lemma.** For all  $k, l \geq 0$ ,  $M_k M_l \leq M_{k+l}$ .

Indeed, for any pair of meanders of orders  $k$  and  $l$  it is possible to define their “sewing”, a meander of order  $k + l$  (see Figure 4.2). Different pairs produce different meanders.

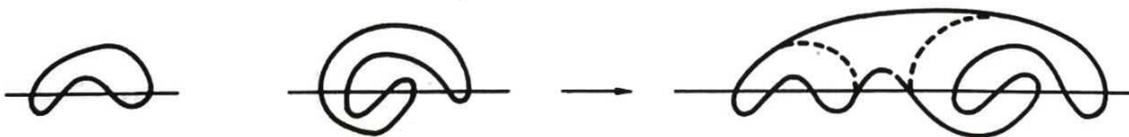


Figure 4.2.

**Proof of the proposition.** Let  $M_k \geq a^k$  for some  $a > 0$ ,  $k \in \mathbb{N}$ . Then, because of the preceding lemma,  $M_{kl} \geq a^{kl}$  for all  $l = 1, 2, \dots$ . Therefore, in view of Proposition 4.1,  $M_n \geq a^{-k} a^n$  for all  $n \geq k$ .

Now assume  $A_M = \sup\{a: \exists k, M_k \geq a^k\}$ . Because of Proposition 4.2 the supremum does exist. Obviously, this value  $A_M$  satisfies the conditions of the proposition.

### 5. Irreducible systems of meanders

Let us encode systems of meanders by words in the alphabet  $\{a, b, c, d\}$  (see 2.2).

**5.1. Definition.** A system of meanders  $W$  is called *irreducible* if no subword of  $W$  forms a system of meanders. The corresponding word will be also called irreducible.

Note that two irreducible words either do not intersect or coincide.

#### Examples

- (a) Any meander forms an irreducible system of meanders.
- (b) The system  $abaccdbd$  (Figure 5.1a) is irreducible. The system  $abacbdcd$  (Figure 5.1b) is reducible since it contains the subword  $acbd$ .



Figure 5.1.

**Notation.** The number of irreducible systems of meanders passing through  $2n$  points is denoted by  $N_n$ . The generating function for the sequence  $N_n$ ,  $n = 0, 1, 2, \dots$ , is denoted by  $N(x)$ :  $N(x) = \sum_{n=0}^{\infty} N_n x^n$ .

#### 5.2. Table

These are the values of  $N_n$  for  $n = 0, 1, \dots, 10$ .

$n$	0	1	2	3	4	5	6	7	8	9	10
$N_n$	1	1	2	8	46	322	2546	21870	199494	1904624	18846714

As we shall see, there is a simple algorithm for computing  $N_n$  (in contrast to  $M_n$ ).

#### 5.3. Admissible subsets

Any system of meanders is unambiguously divided into irreducible components. This geometrically obvious fact nevertheless requires a cumbersome formalization procedure, which this and the following subsections are devoted to.

Let a word  $w = \alpha_1 \alpha_2 \cdots \alpha_{2n}$  in the alphabet  $\{a, b, c, d\}$  represent a system  $W$  of meanders. We call a subset  $A \subset \{1, 2, \dots, 2n\}$  *admissible for the system  $W$*  if it contains together with the left end of each arc its right end, and vice versa. Obviously, for any admissible subset  $A$  it is possible to define the restriction of the system  $W$  onto the set  $A$ , the restriction itself forming a system of meanders given by a subsequence  $\alpha_{i_1} \cdots \alpha_{i_{2k}}$  of letters,  $A = \{i_1, \dots, i_{2k}\}$ ,  $i_1 < \cdots < i_{2k}$ . Obviously, the intersection and union of two admissible subsets are also admissible.

In what follows we will allow ourselves the liberty of being somewhat less than rigorous and identify a subset  $A$  with the corresponding subsequence of letters.

We will call a subword  $\alpha_k \alpha_{k+1} \cdots \alpha_{k+m}$  of a word  $W$  admissible if the subset  $\{k, k+1, \dots, k+m\} \subset \{1, 2, \dots, 2n\}$  is admissible. The intersection of an admissible subset  $A$  and an admissible subword is an admissible subword in  $A$ .

#### 5.4. Splitting of a system of meanders into irreducible components

**Definition.** A subset  $A$  which is admissible for a system  $W$  of meanders is called an *irreducible component* of the system  $W$  if the restriction of  $W$  on  $A$  is irreducible and no other admissible subset  $A_1 \supset A$  has this property.

**Proposition.** *Two irreducible components either do not intersect or coincide.*

**Proof.** Indeed, let  $A_1$  and  $A_2$  be two different irreducible components of the system  $W$  and let neither contain the other. Assume that the intersection  $A_1 \cap A_2$  is not empty. The set  $A_1 \cup A_2$  is admissible and reducible. This means that there exists an admissible subword  $W' \subset A_1 \cup A_2$ . Then  $W'_1 = W' \cap A_1$  is a nonempty admissible subword in  $A_1$ , which contradicts the irreducibility of  $A_1$ .

**Corollary.** *Any system  $W$  of meanders splits unambiguously (modulo a permutation) into irreducible nonintersecting components  $A_1 \cup A_2 \cup \cdots \cup A_k$ .*

This is simple: select the extreme left irreducible subword, delete it, and repeat the process.

**5.5. Description of the language of systems of meanders**

Let  $\lambda$  be the empty word. We introduce an auxiliary symbol  $s$  and represent the following infinite system of substitutions (or “productions”):

$$\begin{aligned}
 s &\rightarrow \lambda \quad (\text{the empty word}) \\
 s &\rightarrow asds \\
 s &\rightarrow asbscsds \\
 &\dots\dots\dots \\
 s &\rightarrow \alpha_1 s \alpha_2 s \dots \alpha_{2n} s \\
 &\dots\dots\dots,
 \end{aligned}
 \tag{5.1}$$

where  $\alpha_1 \alpha_2 \dots \alpha_{2n}$  is an arbitrary irreducible word.

A word in the alphabet  $\{a, b, c, d\}$  is called *deducible* if it can be obtained by a sequential application of the substitutions (5.1). We say that the deduction is *unambiguous* if each deducible system is uniquely representable as a right-hand side of (5.1), where each occurrence of the symbol  $s$  is replaced by a deducible word.

**Proposition.** *The set of deducible words coincides with the set of admissible words, i.e. words describing systems of meanders. The deduction is unambiguous.*

Proposition 5.5 is an immediate corollary of the uniqueness of the decomposition of a system of meanders into irreducible components.

**Note.** We intentionally used a form traditional for describing context-free languages (see [17]). The only difference is that in context-free languages the number of productions is finite. Nevertheless, the analogy is preserved hereafter.

**5.6. Generating noncommutative power series**

Let us introduce a formal power series in noncommuting variables  $a, b, c, d$ :

$$\mathbf{B} = \mathbf{B}(a, b, c, d) = 1 + ad + abcd + acbd + abcbed + \dots,$$

where the right-hand side represents the sum of all admissible words.

Proposition 5.5 admits the following reformulation:

**Proposition.** *The series  $\mathbf{B}$  satisfies the equation*

$$\mathbf{B} = 1 + a\mathbf{B}d\mathbf{B} + a\mathbf{B}b\mathbf{B}c\mathbf{B}d\mathbf{B} + \dots + \alpha_1 \mathbf{B} \alpha_2 \mathbf{B} \dots \alpha_{2n} \mathbf{B} + \dots, \tag{5.2}$$

where the right-hand side contains the sum over all irreducible words  $\alpha_1 \alpha_2 \dots \alpha_{2n}$ .

### 5.7. Functional equation for generating functions

Let  $B(x)$  be the generating function for the number of systems of meanders, i.e. for the squares of Catalan numbers:

$$B(x) = \sum_{n=0}^{\infty} (\text{Cat}_n)^2 x^n.$$

**Theorem.** *The functions  $B(x)$  and  $N(x)$  satisfy the functional equation*

$$B(x) = N(xB^2(x)). \quad (5.3)$$

**Proof.** Replace each of the letters  $a, b, c, d$  in equality (5.2) by the same formal variable  $t$ . We obtain the relation

$$\begin{aligned} B(t^2) &= 1 + t^2 B^2(t^2) + 2t^4 B^4(t^2) + 8t^6 B^6(t^2) + 46t^8 B^8(t^2) \\ &\quad + \cdots + N_n t^{2n} B^{2n}(t^2) + \cdots = N(t^2 B^2(t^2)). \end{aligned}$$

Replacing  $t^2$  by  $x$  we obtain (5.3).

### 5.8. Notes

- (1) Theorem 5.7 can be generalized. Let a set of words have the following property: Between any two letters and at the end of any word belonging to the set, any word from the set may be inserted, and the result will again be a word belonging to the same set. Let  $D(x) = \sum_{n=0}^{\infty} D_n x^n$  be the generating function for the number of words of length  $n$  in the set. A word is irreducible if it does not contain a nonempty subword from the same set. Let  $K(x) = \sum_{n=0}^{\infty} K_n x^n$  be the generating function for the number of irreducible words. Then the equality

$$D(x) = K(xD^2(x)) \quad (5.4)$$

holds. The theorem was proved in this form in [19].

- (2) The above method of proof is borrowed from the theory of formal languages (see [16], [17]) and goes back to the papers by Schützenberger (see [18]). In this theory the number of productions is finite, as a rule, and therefore the equations are algebraic. Languages with an infinite number of productions are examined in [3].
- (3) As a very simple illustration of applying the same method, consider the Dick language (see 2.2). Let  $x$  denote the left bracket and  $y$  the right one. There are only two productions; here they also correspond to irreducible words  $\lambda$  and  $xy$ :

$$\begin{aligned} s &\rightarrow \lambda \\ s &\rightarrow xsys. \end{aligned}$$

For the generating function of Catalan numbers we obtain the equation

$$\text{Cat}(t^2) = 1 + t^2 \text{Cat}^2(t^2) \quad \text{or} \quad \text{Cat}(x) = 1 + x \text{Cat}^2(x). \quad (5.5)$$

The function  $1 + x$  acts here as the function  $K$ .

- (4) Equality (5.4) yields an efficient algorithm for calculating the numbers  $K_n$ . If we know the first  $n$  coefficients of the function  $D$  (in our case, the squares of Catalan numbers) we can find the first  $n$  coefficients of the function  $K(x)$ .

### 5.9. Generating function for squares of Catalan numbers

**Lemma.** *Let  $f(x) = a_0 + a_1x + a_2x^2 + \dots$ , the generating function for a sequence  $a_0, a_1, \dots$ , be analytical in a neighborhood of zero in  $\mathbb{C}$ . Then the generating function  $F(x)$  for the sequence of squares  $a_0^2, a_1^2, \dots$ ,*

$$F(x) = a_0^2 + a_1^2x + a_2^2x^2 + \dots,$$

*is analytical in a neighborhood of zero and representable as*

$$F(x) = F(t^2) = \text{Res}_{\lambda=0} \frac{f(\lambda t)f(\lambda^{-1}t)}{\lambda}. \quad (5.6)$$

**Proof.** The analytical nature of  $F(x)$  follows from the obvious estimates of the convergence radius. Multiplying the series for  $f(\lambda t)$  and  $f(\lambda^{-1}t)$  we see that the coefficients in front of the monomials not containing powers of  $\lambda$  are exactly  $a_0^2, a_1^2, \dots$ .

This lemma was in fact proved by Parseval in 1805 (see [14]).

It remains to substitute into (5.6) the generating function for Catalan numbers.

**Proposition.**

$$B(t^2) = \frac{1}{4t^2} \left( -1 + \frac{1}{2\pi} \int_0^{2\pi} \sqrt{1 - 8t \cos \phi + 16t^2} d\phi \right). \quad (5.7)$$

**Proof.** As follows from (5.5), the generating function for Catalan numbers equals

$$\text{Cat}(x) = \frac{1 - \sqrt{1 - 4x}}{2x}.$$

We obtain

$$\begin{aligned}
 B(t^2) &= \operatorname{Res}_{\lambda=0} \frac{\operatorname{Cat}(\lambda t) \operatorname{Cat}(\lambda^{-1}t)}{\lambda} \\
 &= \operatorname{Res}_{\lambda=0} \frac{1}{\lambda} \frac{1 - \sqrt{1 - 4\lambda t}}{2\lambda t} \frac{1 - \sqrt{1 - 4\lambda^{-1}t}}{2\lambda^{-1}t} \\
 &= \frac{1}{4t^2} \operatorname{Res}_{\lambda=0} \frac{1}{\lambda} (1 - \sqrt{1 - 4\lambda t})(1 - \sqrt{1 - 4\lambda^{-1}t}) \\
 &= \frac{1}{4t^2} \left[ \operatorname{Res}_{\lambda=0} \frac{1}{\lambda} (1 - \sqrt{1 - 4\lambda t}) - \operatorname{Res}_{\lambda=0} \frac{1}{\lambda} \sqrt{1 - 4\lambda^{-1}t} \right. \\
 &\quad \left. + \operatorname{Res}_{\lambda=0} \frac{1}{\lambda} \sqrt{1 - 4\lambda t - 4\lambda^{-1}t + 16t^2} \right].
 \end{aligned}$$

Decomposing the expression under the residue symbols into powers of  $\lambda$  we check that the first residue equals 0 and the second one equals 1. Representing the last residue as an integral along the circle  $|\lambda| = 1$ , we obtain formula (5.7).

**5.10. Convergence radius of generating series for irreducible systems of meanders**

**Lemma.** *Let the power series  $D(x)$  and  $K(x)$  with nonnegative coefficients be related by (5.4) and let  $r > 0$  be the radius of convergence of the series  $D(x)$ , and let the numerical series  $D(r)$  converge. Then the radius of convergence of the series  $K(x)$  is not smaller than*

$$\rho = rD^2(r). \tag{5.8}$$

**Proof.** Let us demonstrate that the series  $K(x)$  converges absolutely at any point  $x$ ,  $|x| = q < \rho$ . Since the function  $x D^2(x)$  is monotonous and continuous on the segment  $[0, r]$ , there exists  $p \in [0, r]$  such that  $p D^2(p) = q$ . Denote by  $K^{(n)}(x) = K_0 + K_1 x + \dots + K_n x^n$  the partial sum of the series  $K(x)$ . Then

$$|K^{(n)}(x)| \leq K^{(n)}(q) = K^{(n)}(p D^2(p)) \leq D(p);$$

the latter inequality follows from the fact that all the coefficients of the series  $K^{(n)}(p D^2(p))$  do not exceed the corresponding coefficients of the series  $K(p D^2(p)) = D(p)$  due to the nonnegativity of the coefficients of the series  $K(x)$ . The lemma is proved.

The example of Equation (5.5) for the generating function of Catalan numbers shows that the radius of convergence of the function  $K(x)$  may,

in general, be greater than  $\rho$ . Indeed, the radius of convergence of  $\text{Cat}(x)$  equals  $r = \frac{1}{4}$ ,  $\text{Cat}(r) = 2$ , so that  $\rho = 1$ . And the radius of convergence of the series  $K(x) = 1 + x$  is infinite.

**Proposition.** *The radius of convergence of the series  $B(x)$  is equal to  $r = \frac{1}{16}$ . The series converges at the point  $\frac{1}{16}$  and  $B(\frac{1}{16}) = 4(4/\pi - 1)$ . Finally,*

$$\rho = rB^2(r) = \frac{1}{16}B^2(\frac{1}{16}) = \left(\frac{4 - \pi}{\pi}\right)^2. \quad (5.9)$$

**Proof.** The value  $r = \frac{1}{16}$ , as well as the convergence of the series  $B(x)$  at the point  $\frac{1}{16}$ , follows from formula (2.2).

To the value  $r = \frac{1}{16}$  there corresponds the value  $t = \frac{1}{4}$  in formula (5.7). The integral, as a function of the parameter  $t$ , is continuous on the segment  $[0, \frac{1}{4}]$ . Its calculation at the point  $t = \frac{1}{4}$  is reduced to an elementary integral.

### 5.11. Radius of convergence. Continuation

We have proved that the radius of convergence of the series  $N(x)$  is at least  $\rho = ((4 - \pi)/\pi)^2$ . In fact, it is equal to  $((4 - \pi)/\pi)^2$ .

**Lemma.** *Let the conditions of Lemma 5.10 be satisfied and let the series  $D'(r)$  converge. Then the radius of convergence of the series  $K(x)$  equals  $\rho$ .*

**Proof.** Let us rewrite equality (5.4) in the form

$$D(t^2) = K(t^2D^2(t^2)). \quad (5.10)$$

Setting  $\lambda = \phi(t) = tD(t^2)$  and multiplying both sides of equality (5.10) by  $t$ , we obtain

$$\phi(t) = tK(\phi^2(t)) = t\omega(\phi(t)),$$

where  $\omega(\lambda) = K(\lambda^2)$ . Hence

$$\omega(\lambda) = \frac{\lambda}{\phi^{-1}(\lambda)}. \quad (5.11)$$

The functions  $\omega(\lambda)$  and  $\phi^{-1}(\lambda)$  are defined and holomorphic in the circle of radius  $\sqrt{\rho}$ . Our lemma will be proved if we demonstrate that  $\phi^{-1}(\lambda)$  cannot be extended holomorphically in any neighborhood of the point  $\sqrt{\rho}$ .

Let  $\phi^{-1}(\lambda)$  have a holomorphic extension in a neighborhood of the point  $\sqrt{\rho}$ . Then

$$(\phi^{-1})'(\sqrt{\rho}) = \lim_{\lambda \rightarrow \sqrt{\rho}-0} (\phi^{-1})'(\lambda) = \frac{1}{\lim_{t \rightarrow r-0} \phi'(t)}.$$

The latter limit exists and is positive due to the conditions of the lemma. This means that the function  $\phi^{-1}$  is invertible in a neighborhood of the point  $\sqrt{\rho}$ , which contradicts the conditions of Lemma 5.10.

**Proposition.** *The radius of convergence of the series  $N(x)$  is equal to*

$$((4 - \pi)/\pi)^2.$$

To check this, it suffices to differentiate term by term the power series for  $\phi$  and use the fact that, according to formula (2.2), the coefficients of the function  $B(t^2)$  increase as  $\text{const} \cdot 4^n \cdot n^{-3}$ .

**Notes.** (1) Similar arguments for Catalan numbers are not valid: The series for the derivative diverges at the point  $t = \frac{1}{2}$  due to the exponent  $-\frac{3}{2}$  in the asymptotics  $\text{const} \cdot 2^n \cdot n^{-3/2}$ .

(2) Replacing  $k' = (1 + 4t)/(1 - 4t)$ ,  $k^2 = 1 - k'^2$ , in the integral of formula (5.7) we reduce it to a complete elliptic integral of genus II. The singular point  $t = \frac{1}{4}$  corresponds to the singular point  $k = \infty$  of the elliptic integral.

Propositions 5.10 and 5.11 together prove Theorem 2.2.

## 6. Empirical estimates

### 6.1. General remarks

When enumerating combinatorial objects, numerical sequences often emerge having asymptotics of the form

$$a_n \sim \text{const} \cdot A^n \cdot n^{-\alpha} \quad (6.1)$$

(the parameter  $\alpha$  often turns out to be an integer or a half-integer). Examples can be found in [2], [6].

There exist at least two large classes for which such asymptotics can be validated theoretically. The first class corresponds to the generating functions with a unique singularity on the boundary of the disk of convergence, under certain conditions on the character of the singularity (see [6], Section 9.5). The second important class corresponds to the case in which the fraction

$a_{n+1}/a_n$  is a rational function of  $n$ :

$$\frac{a_{n+1}}{a_n} = \frac{P(n)}{Q(n)}, \quad (6.2)$$

where  $P$  and  $Q$  are polynomials of the same degree. The generating functions for all such sequences  $a_n$  are generalized hypergeometric functions.

Formula (6.2) leads us to think it reasonable, in studying a sequence whose expected asymptotics has the form (6.1), to consider the function  $f(n) = a_{n+1}/a_n$  and to construct (multipoint) Padé approximations for it. The present section is devoted to the application of this approach to meandric numbers.

## 6.2. Basic lemma

**Lemma.** *Let a sequence of positive integers  $a_n$  be such that*

$$\frac{a_{n+1}}{a_n} = A \frac{n^k + a_1 n^{k-1} + \cdots + a_k}{n^k + b_1 n^{k-1} + \cdots + b_k}, \quad (6.3)$$

and  $a_1 \neq b_1$ . Then

$$a_n \sim \text{const} \cdot A^n \cdot n^{-(b_1 - a_1)}. \quad (6.4)$$

**Example.** For Catalan numbers

$$\frac{\text{Cat}_{n+1}}{\text{Cat}_n} = \frac{4n+2}{n+2} = 4 \frac{n + \frac{1}{2}}{n+2},$$

whence  $\text{Cat}_n \sim \text{const} \cdot 4^n \cdot n^{-\frac{3}{2}}$ , which is in accord with formula (2.2).

**Proof of the lemma.** Represent the right-hand side of formula (6.3) as a series in powers of  $1/n$ :

$$\frac{a_{n+1}}{a_n} = A \left( 1 - \frac{b_1 - a_1}{n} + \frac{\alpha_2}{n^2} + \cdots \right) = A \left( 1 - \frac{\alpha}{n} + O\left(\frac{1}{n^2}\right) \right).$$

Taking the logarithm of the latter equality for  $n = 1, 2, \dots$  and summing up the result term by term, we obtain

$$\ln a_{n+1} - \ln a_1 \approx n \ln A - \alpha \left( 1 + \frac{1}{2} + \cdots + \frac{1}{n} \right) \approx n \ln A - \alpha \ln n,$$

which yields the asymptotics (6.4). We omit the standard estimates of the residual terms.

### 6.3. Continued fractions

Consider a function of an integer argument:

$$f(n) = \frac{M_{n+1}}{M_n}, \quad n = 1, 2, \dots$$

We know, from the table in 3.3, the first 13 values of this function:  $f(1) = \frac{2}{1} = 2$ ,  $f(2) = \frac{8}{2} = 4$ ,  $f(3) = \frac{42}{8} = 5\frac{1}{4}$ , etc. We will look for the rational function which takes the same values. It is convenient to look for the answer in the form of a continued fraction:

$$f(n) = f_1 + \frac{n-1}{f_2 + \frac{n-2}{f_3 + \frac{n-3}{f_4 + \dots}}}; \quad (6.5)$$

this form is convenient because when new values of the function  $f(n)$  emerge we do not have to calculate anew the values  $f_1, f_2, \dots$  obtained earlier. When  $f(n)$  is representable as  $P(n)/Q(n)$ , as it is for Catalan numbers or their squares, the continued fraction (6.5) terminates.

We actually made the computations not only on the basis of the segment  $[1, 13]$  but also on all its subsegments with an odd number of points.

### 6.4. Results of the computations

These are the values of the constants  $A_M$  and  $\alpha$  computed for the sequence  $M_n$  by using 14 points (i.e. 13 relations  $M_{n+1}/M_n$ ):

$$A_M = 12.26239 \dots,$$

$$\alpha_M = 3.41998 \dots$$

We would like to estimate the accuracy of the results by testing the method on another sequence with a similar conduct. The Catalan numbers (and their squares) are not suitable for this purpose since the method yields an accurate result for them due to the above considerations. Therefore we will apply the method to the sequence of irreducible systems  $N_n$ , again taking the first 14 elements of the sequence. This is the result:

$$A_N = 13.3923 \dots,$$

$$\alpha_N = 2.9685 \dots,$$

while

$$\left(\frac{\pi}{4-\pi}\right)^2 = 13.39408 \dots$$

Thus, the approximate value of the constant  $A_N$  coincides with the actual value up to the fourth significant digit. The character of the behaviour when approaching the constant  $\alpha$  testifies to the fact that in this case the accuracy of the approximation will be much lower. However, there are grounds to believe that the above constant in the cases considered is either an integer or a half-integer. This allows us to formulate the following hypotheses:

**Hypothesis 1.**  $M_n \sim \text{const} \cdot A_M \cdot n^{-\alpha_M}$ , where  $A_M = 12.26 \dots$ ,  $\alpha_M = \frac{7}{2}$ .

**Hypothesis 2.**  $N_n \sim \text{const} \cdot (\pi/(4 - \pi))^{2n} \cdot n^{-\alpha_N}$ , where  $\alpha_N = 3$ .

## 7. Method of matrix integrals

The method of matrix integrals was invented by physicists (see, for example, [2] and the references therein) and reinvented independently by mathematicians [7]. It allows one to solve some mathematical problems not yielding to other methods. It has not yet provided an answer to the problem of meanders, as we have not succeeded in calculating the emerging matrix integral. Nevertheless, we decided to include the material in this paper for two reasons. First, the relations between the problem of meanders and quantum field theory, as well as the method of matrix integrals, are of independent interest. Second, the problem of calculating the integral may be solved by a reader of the present paper.

The result presented here belongs to Vladimir Kazakov and Ivan Kostov ([8]) but has not been published. The authors are grateful for the permission to include it in the present paper.

### 7.1. Basic information on Gaussian measures

Let  $B$  be a positive definite matrix giving a quadratic form in  $\mathbf{R}^n$ . Let us introduce a *Gaussian measure* on  $\mathbf{R}^n$ :

$$d\mu(x) = (2\pi)^{-n/2} (\det B)^{1/2} \exp\left\{-\frac{1}{2}(Bx, x)\right\} dx. \quad (7.1)$$

The factor in front of the exponent normalizes the measure so that the measure of the entire space  $\mathbf{R}^n$  is equal to 1. For  $f = f(x)$ ,  $x \in \mathbf{R}^n$ , we borrow from physics the notation for the average value of  $f$ :

$$\langle f \rangle = \int_{\mathbf{R}^n} f(x) d\mu(x).$$

**Lemma.**  $\langle x_i x_j \rangle = a_{ij}$ , where  $A = (a_{ij}) = B^{-1}$ .

Using this lemma it is easy to obtain the average values  $\langle fg \rangle$ , where  $f$  and  $g$  are linear functions. The average value of any monomial of odd degree equals zero. The average values of monomials of even degree are obtained by means of the following theorem.

**Theorem.** *Let  $f_1, f_2, \dots, f_{2m}$  be linear functions of  $x$ , not necessarily different. Then*

$$\langle f_1 f_2 \cdots f_{2m} \rangle = \sum \langle f_{p_1} f_{q_1} \rangle \langle f_{p_2} f_{q_2} \rangle \cdots \langle f_{p_m} f_{q_m} \rangle, \tag{7.2}$$

where the sum is taken over all  $(2m - 1)!!$  sets of indices  $p_1 < q_1, \dots, p_m < q_m, p_1 < p_2 < \cdots < p_m$ , i.e. over all partitions of the set  $\{1, 2, \dots, 2m\}$  into pairs.

Formula (7.2) is known in physics as Wick's formula. It is a highly special case of the formula, well known in probability theory, expressing the moments of a probabilistic distribution via its semi-invariants ([11]) or, in a still more general form, of the formula connecting the Taylor series of a function of many variables and of its logarithm.

**7.2. Gaussian measure on Hermitian matrices**

The following construction is a special case of that in 7.1.

Let  $\mathbf{H}_N$  be the space of Hermitian matrices  $H = (h_{ij})$  of size  $N \times N$ . We introduce the coordinates  $x_{ij} = \text{Re } h_{ij}, i \leq j, y_{ij} = \text{Im } h_{ij}, i < j$ , on it. In these coordinates the space  $\mathbf{H}_N$  is isomorphic to an  $N^2$ -dimensional Euclidean space. Denote by  $dv(H)$  the usual Lebesgue measure in this space:

$$dv(H) = \prod_{i=1}^N dx_{ii} \prod_{i < j} dx_{ij} dy_{ij}.$$

In order to define a Gaussian measure on  $\mathbf{H}_N$ , it is necessary to pick a positive definite quadratic form on  $\mathbf{H}_N$ . We take  $\text{tr } H^2$  as this form. Obviously,

$$\text{tr } H^2 = \sum_{i,j=1}^N h_{ij} h_{ji} = \sum_{i,j=1}^N h_{ij} \bar{h}_{ij} = \sum_{i=1}^N x_{ii}^2 + 2 \sum_{i < j} (x_{ij}^2 + y_{ij}^2). \tag{7.3}$$

The matrix of this quadratic form has dimension  $N^2 \times N^2$ . Formula (7.3) states that the matrix is diagonal in the coordinates  $x_{ij}, y_{ij}$ ; it has  $N$  1's (corresponding to the coordinates  $x_{ii}$ ) and  $N^2 - N$  2's (corresponding to the coordinates  $x_{ij}, y_{ij}, 1 \leq i < j \leq N$ ). Its determinant equals  $2^{N^2 - N}$ . In accordance with formula (7.1), we can define a Gaussian measure on  $\mathbf{H}_N$ :

$$d\mu(H) = (2\pi)^{-N^2/2} 2^{(N^2 - N)/2} \exp\{-\frac{1}{2} \text{tr } H^2\} dv(H). \tag{7.4}$$

We take the matrix entries  $h_{ij}$  as linear functions of  $x_{ij}, y_{ij}$ .

**Lemma.**  $\langle h_{ij} \cdot h_{ji} \rangle = 1$ ; if  $(i, j) \neq (l, k)$ , then  $\langle h_{ij} \cdot h_{kl} \rangle = 0$ .

**Proof.** The matrix inverse of the matrix of the quadratic form  $\text{tr } H^2$  is diagonal and contains on the diagonal  $N$  1's and  $N^2 - N$   $\frac{1}{2}$ 's. Therefore

$$\begin{aligned} \langle h_{ii}^2 \rangle &= \langle x_{ii}^2 \rangle = 1 \quad \text{when } i = j; \\ \langle h_{ij} h_{ji} \rangle &= \langle h_{ij} \bar{h}_{ij} \rangle = \langle x_{ij}^2 + y_{ij}^2 \rangle = \frac{1}{2} + \frac{1}{2} = 1 \quad \text{when } i < j; \end{aligned}$$

we obviously get zero in the remaining cases.

### 7.3. The basic integral

Let  $(\mathbf{H}_N)^{2q}$  be the space of sets  $(H_1, \dots, H_q, G_1, \dots, G_q)$  of Hermitian matrices of dimension  $N \times N$ . Take the Gaussian measure on  $(\mathbf{H}_N)^{2q}$  equal to the product of the measures (7.4). The following function will play the main role in the constructions that follow:

$$\begin{aligned} Z(s, q, N) &= \frac{1}{N^{2q}} \left\langle \exp \left\{ -\frac{s}{N} \text{tr} \sum_{k,l=1}^q H_k G_l H_k G_l \right\} \right\rangle \\ &= \frac{1}{N^{2q}} \int_{(\mathbf{H}_N)^{2q}} \exp \left\{ -\frac{s}{N} \text{tr} \sum_{k,l=1}^q H_k G_l H_k G_l \right\} \prod_{k=1}^q d\mu(H_k) \prod_{l=1}^q d\mu(G_l). \end{aligned} \tag{7.5}$$

The function  $Z$  can be regarded as a partition function of a quantum field theory model. In this model the space consists of  $2q$  isolated points; the field defined on this space takes values in the space  $\mathbf{H}_N$  of Hermitian matrices; the Gaussian measure on  $(\mathbf{H}_N)^{2q}$  corresponds to the free field; the polynomial of degree four in the exponent is the Lagrangian of the interaction; the parameter  $s$  plays the role of cosmological constant, and  $q$  and  $N$  are two other parameters of the theory.

Further, the arguments are nothing but a version of the method of perturbation theory; meanders emerge as a subclass of Feynman diagrams.

### 7.4. A perturbation theory series and Feynman diagrams

Let us expand the exponent under the integral (7.5) into the series

$$\exp \left\{ -\frac{s}{N} \text{tr} \sum_{k,l=1}^q H_k G_l H_k G_l \right\} = \sum_{m \geq 0} \frac{(-1)^m s^m}{m! N^m} \left( \text{tr} \sum_{k,l=1}^q H_k G_l H_k G_l \right)^m. \tag{7.6}$$

The expression in the parantheses may be written in greater detail:

$$\text{tr} \sum_{k,l=1}^q H_k G_l H_k G_l = \sum_{k,l=1}^q \sum_{i_1, i_2, i_3, i_4=1}^N h_{i_1 i_2}^{(k)} g_{i_2 i_3}^{(l)} h_{i_3 i_4}^{(k)} g_{i_4 i_1}^{(l)}, \tag{7.7}$$

where  $h^{(k)}$  and  $g^{(l)}$  are the entries of the matrices  $H_k$  and  $G_l$ , respectively. Raising the sum to degree  $m$  we get

$$\sum_{k_1 l_1 = 1}^q \cdots \sum_{k_m l_m = 1}^q \sum_{i_1^1 i_2^1 i_3^1 i_4^1 = 1}^N \cdots \sum_{i_1^m i_2^m i_3^m i_4^m = 1}^N h_{i_1^1 i_2^1}^{(k_1)} g_{i_2^1 i_3^1}^{(l_1)} h_{i_3^1 i_4^1}^{(k_1)} g_{i_4^1 i_1^1}^{(l_1)} \cdots h_{i_3^m i_4^m}^{(k_m)} g_{i_4^m i_1^m}^{(l_m)}. \quad (7.8)$$

These expressions are too cumbersome; but Feynman diagrams were invented for the special purpose of avoiding indices.

Expressions (7.7) may be illustrated by the following geometrical image:

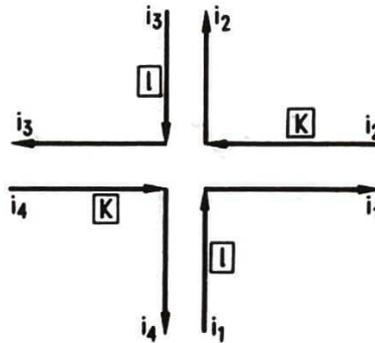


Figure 7.1.

It reminds us of a crossroad; along the roads are the indices  $i_1, i_2, i_3, i_4$ ; an incoming arrow corresponds to the first index of a matrix entry, an outgoing arrow corresponds to the second one. The roads are colored in colors  $k$  and  $l$ . (There are  $2q$  colors in all. But the first road is painted in a color from the first half, and the second in one from the second half; this also corresponds to the positions of the matrices  $H$  and  $G$  in (7.7).) To expression (7.7) raised to degree  $m$  there corresponds a set of  $m$  copies of such “crossroads” labelled by the numbers  $1, 2, \dots, m$ .

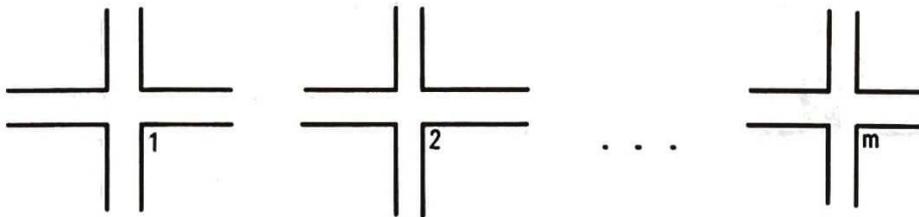


Figure 7.2.

Now we apply Wick’s formula to each product of the factors of the sum (7.8). It follows from the lemma in 7.2 that in the majority of cases the result will be

zero (it suffices if at least one pairing yields zero for this result). The result is nonzero (it is then equal to 1) if and only if

- (a) in each pairing, the participating matrix entries have the same color (i.e. belong to the same matrix);
- (b) the conditions on the indices indicated in the lemma in 7.2 are fulfilled.

The pairing will be represented graphically as in Figure 7.3.

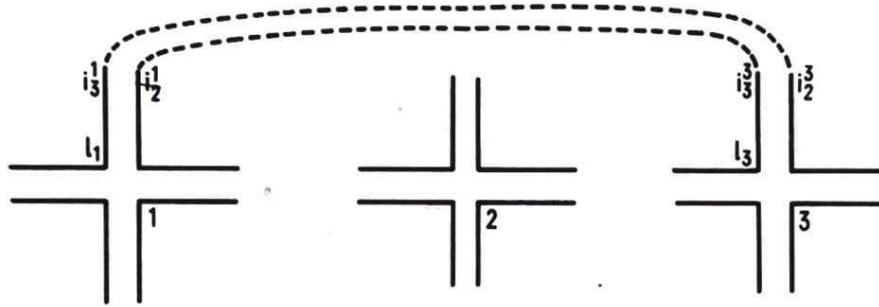


Figure 7.3.

In this figure a pair consists of  $g_{i_2^1 i_3^1}^{(1)}$  and  $g_{i_2^3 i_3^3}^{(3)}$ . To get a nonzero result, it is necessary and sufficient that the equalities  $l_1 = l_3$ ,  $i_2^1 = i_2^3$ ,  $i_3^1 = i_3^3$  be satisfied. The directions of the arrows turn out to be consistent.

Partitioning all the factors in (7.8) into pairs leads to all of the roads leaving the crossroads being “sewn” pairwise (see Figure 7.4). The arrows corresponding to the matrix indices are sewn into cycles. Let  $p$  be the number of these cycles. Then a nonzero result is obtained for exactly  $N^p$  sets of the  $N^{4m}$  possible sets of indices  $i_1^1, \dots, i_4^m$ , since the indices appearing in a cycle must all be equal to one another.

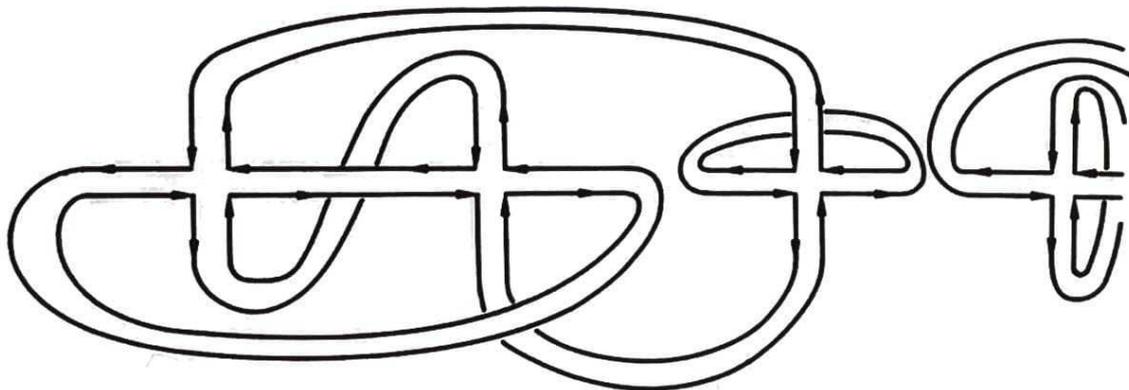


Figure 7.4.

Further, all the roads are divided into some number of one-colored circles. Note that the circles painted by colors from the same half of the complete list of colors do not intersect (see Figure 7.1). Let  $r$  be the number of all circles. Then there exist  $q^r$  possible colorings of the circles.

Thus if we take the average of the sum (7.8) and apply to it Wick's formula, each graph of the type drawn in Figure 7.4 contributes  $N^p \cdot q^r$  to the total sum, where  $p$  is the number of cycles formed by the arrows and  $r$  is the number of circular roads. The formal series for the function  $Z$  defined in (7.5), taking (7.6) into account, is

$$Z(s, q, N) = \frac{1}{N^2} \sum \frac{(-1)^m s^m}{m! N^m} \cdot N^p q^r, \quad (7.9)$$

where the sum is taken over all graphs having  $m$  vertices,  $p$  cycles of arrows and  $r$  circular roads. We label the vertices of the graphs and label the germs of edges starting at each vertex.

### 7.5. Euler characteristic

Let us rewrite formula (7.9) in the form

$$Z(s, q, N) = \sum_{m=0}^{\infty} \frac{(-1)^m s^m}{m! N^{m+2}} \cdot P_m(q, N), \quad (7.10)$$

where  $P_m(q, N)$  is a polynomial in the variables  $q$  and  $N$  whose lowest degree in  $q$  equals two. Denote the coefficient at  $q^2$  by  $Q_m(N)$ ; this is a polynomial of  $N$ . Thus the series

$$Y(s, N) = \sum_{m=0}^{\infty} \frac{(-1)^m s^m}{m! N^{m+2}} \cdot Q_m(N), \quad (7.11)$$

which is the "coefficient at  $q^2$ " of the series  $Z$ , is defined correctly.

From the combinatorial point of view, the coefficient at  $q^2$  singles out all pictures that have exactly two circular roads. The corresponding graphs are automatically connected. A connected picture is transformed into a closed oriented two-dimensional manifold if each cycle of arrows is glued by a cell. Let us calculate the Euler characteristic of this manifold. The number of vertices is  $m$ , the number of edges is  $4m/2 = 2m$ , the number of cells is  $p$  (as many as there are cycles of arrows). Thus, the Euler characteristic equals

$$2 - 2g = m - 2m + p = -m + p,$$

whence

$$2g = m + 2 - p, \quad (7.12)$$

where  $g$  is the genus of the manifold. Now formula (7.11) is rewritten in the form:

$$\begin{aligned} Y(s, N) &= \sum \frac{(-1)^m}{m!} \frac{s^m}{N^{m+2}} \cdot N^p = \sum \frac{(-1)^m}{m!} s^m \cdot N^{-2g} \\ &= Y_0(s) + \frac{1}{N^2} Y_1(s) + \frac{1}{N^4} Y_2(s) + \dots, \end{aligned} \quad (7.13)$$

where the sum is taken over all graphs "with two roads".

The discussion of the convergence of the series (7.13) is outside the framework of the present paper. Note only the following feature: The series in the inverse powers of  $N$  is asymptotic (when  $N \rightarrow \infty$ ); in contrast, the functions  $Y_0(s)$ ,  $Y_1(s)$ ,  $\dots$  are analytical in a neighborhood of the point  $s = 0$  but their radii of convergence decrease as the number of function increases.

Obviously, only meanders yield a surface of genus  $g = 0$ , i.e. a sphere. For this reason we are interested in the principal term of the asymptotics in  $N$ , i.e. the function  $Y_0(s)$ .

## 7.6. Final result

Let us repeat the two stages of obtaining the function  $Y_0(s)$  in 7.5:

- (1) isolation of the coefficient  $Y(s, N)$  at  $q^2$  in the function  $Z(s, q, N)$ ;
- (2) isolation of the principal term  $Y_0(s)$  in the asymptotic series in  $1/N$  of the function  $Y(s, N)$ .

**Theorem.** *The coefficient at  $s^{2n}$  in the series  $Y_0(s)$  equals*

$$\frac{2^{2n-2}}{n} M_n, \quad (7.14)$$

where  $M_n$  is the number of meanders of order  $n$ .

**Proof.** Let us count the number of  $2n$  sewn labelled crosses with ends labelled and colored in two colors in the planar graph consisting of two one-colored circles. The edges of one color may be sewn into a circle in  $(4n - 2)!!$  ways. After the sewing is performed, the edges of the second color may be sewn into a circle in  $M_n$  ways. As a result, the coefficient at  $s^{2n}$  turns out to be equal to

$$\frac{(-1)^{2n}}{(2n)!} (4n - 2)!! M_n = \frac{2^{2n-1}(2n - 1)!}{(2n)!} M_n = \frac{2^{2n-2}}{n} M_n, \quad \text{Q.E.D.}$$

**Note.** The series  $Y_0(s)$  does not contain terms with odd degrees.

Having differentiated the series  $Y_0(s)$  we obtain

**Corollary.**  $2sY'_0(s) = M(4s^2)$ , where  $M(s)$  is the generating function for meandric numbers.

**7.7. On the computation of matrix integrals**

As noted above, the integral (7.5) has not yet been computed, so we have no final solution to the problem of meanders. In the physical literature, a vast store of techniques has been accumulated for computing such integrals (see [2]). However, all of them allow one to compute a comparatively narrow class of integrals. The development of techniques for computing matrix integrals is of special interest for physics. (In physics what is often meant by “computing an integral” is obtaining the first terms of the asymptotics.)

There is a technique, however, which is applied practically always and it is worth mentioning. It consists in passing to the so-called “polar coordinates” in the space of Hermitian matrices, i.e. in replacing a matrix  $H \in \mathbf{H}$  by a pair  $(U, \Lambda)$ , where  $U \in \mathbf{U}(N)$  is a unitary matrix,  $\Lambda$  is a diagonal matrix and

$$H = U\Lambda U^{-1}. \tag{7.15}$$

The representation (7.15) is ambiguous: the matrix  $\Lambda$  is determined to within permutation of the diagonal entries (and the corresponding permutation of the columns of  $U$ ), and the matrix  $U$  (when the characteristic values of  $\Lambda$  are distinct) to within multiplication by a matrix of the form

$$\begin{bmatrix} e^{i\phi_1} & & & 0 \\ & e^{i\phi_2} & & \\ & & \ddots & \\ 0 & & & e^{i\phi_N} \end{bmatrix}.$$

This ambiguity leads to the factor  $1/((2\pi)^N \cdot N!)$  in the following statement.

**Proposition.** ([2]) *Let the function  $f(H)$ ,  $H \in \mathbf{H}$ , be unitary invariant. Then*

$$\int_{\mathbf{H}} f(H) dv(H) = 2^{-(N^2 - N)/2} \frac{1}{(2\pi)^N N!} \text{vol}(\mathbf{U}(N)) \times \int_{\mathbf{R}^n} f(\Lambda) \prod_{i < j} (\lambda_i - \lambda_j)^2 d\lambda_1 \cdots d\lambda_N, \tag{7.16}$$

where  $\text{vol}(\mathbf{U}(N))$  is the volume of the unitary group induced by its natural imbedding into the space of complex matrices of dimension  $N \times N$ .

The volume of the unitary group is

$$\text{vol}(\mathbf{U}(N)) = \frac{(2\pi)^{(N^2 + N)/2}}{\prod_{k=1}^{N-1} k!}.$$

For instance, when

$$f(H) = \exp\left\{-\frac{1}{2} \operatorname{tr} H^2 - \frac{s}{N} \operatorname{tr} H^4\right\}$$

we should obviously substitute the following into (7.16):

$$f(\Lambda) = \exp\left\{-\frac{1}{2} \sum_{i=1}^N \lambda_i^2 - \frac{s}{N} \sum_{i=1}^N \lambda_i^4\right\}.$$

Unfortunately, this method alone does not allow us to compute the integral (7.5).

### References

- [1] V. I. Arnol'd, *The branched covering of  $CP^2 \rightarrow S^4$ , hyperbolicity and projective topology*, Siberian Math. J. **29**:5 (1988), 717–726. (Originally published in *Sibirsk. Math. Zh.* **29**:5 (1988), 36–47.) (MR 90a:57037)
- [2] D. Bessis, C. Itzykson and J. B. Zuber, *Quantum field theory techniques in graphical enumeration*, Adv. in Appl. Math. **1**:2 (1980), 109–157. (MR 83j:81067)
- [3] J. R. Goldman, *Formal languages and enumeration*, J. Combin. Theory Ser. A, **24**:3 (1978), 318–338. (MR 81m:65009a)
- [4] I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, Wiley, NY, 1983. (MR 84m:05002)
- [5] M.-P. Delest and G. Viennot, *Algebraic languages and polyominoes enumeration*, Theor. Comput. Sci. **34**:1 (1984), 169–206. (MR 86e:68062)
- [6] F. Harary and E. Palmer, *Graphical Enumeration*, Academic Press, NY, 1973. (MR 50 # 9682)
- [7] J. L. Harer and D. Zagier, *The Euler characteristic of the moduli space of curves*, Invent. Math. **85**:3 (1986), 457–485. (MR 87i:32031)
- [8] V. A. Kazakov, Private communication, 1988.
- [9] J. E. Koehler, *Folding a strip of stamps*, J. Combin. Theory. **5**:2 (1968), 135–152. (MR 37 # 3945)
- [10] W. F. Lunnon, *A map-folding problem*, Math. Comp. **22** (1968), 193–199. (MR 36 # 5009)
- [11] V. A. Malyshev and R. A. Minos, *Gibbs Random Fields* (in Russian), "Nauka", Moscow, 1985. (MR 87h:82001)
- [12] A. Phillips, Letter to the authors, September 1988.
- [13] A. Phillips, *Simple alternating transit mazes*, preprint, June 1988; Addenda & Errata, September 1988.
- [14] G. Pólya and G. Szegő, *Aufgaben und Lehrsätze aus der Analysis*, Springer, Berlin, 1964. (MR 30 # 1219ab)
- [15] P. Rosenstiehl, *Planar permutations defined by two intersecting Jordan curves*, in *Graph Theory and Combinatorics* (Cambridge, 1983), 259–271, Academic Press, London, 1984. (MR 86f:05006)
- [16] A. Salomaa, *Jewels of Formal Language Theory*, Computer Science Press, Rockville, MD, 1981. (MR 83f:68091)
- [17] A. Salomaa and M. Soittola, *Automata-theoretic Aspects of Formal Power Series*, Springer, NY, 1978. (MR 58 # 3698)
- [18] M.-P. Schützenberger, *Certain elementary families of automata*, in *Proc. Symp. Math. Theory of Automata* (New York, 1962), 139–153, Polytechnic Inst. Brooklyn, Brooklyn, NY, 1962. (MR 29 # 5696)
- [19] T. R. S. Walsh and A. B. Lehman, *Counting rooted maps by genus, III*, J. Combin. Theory. Ser. B **18**:3 (1975), 222–259. (MR 51 # 10154)