PHASIS
Journal Program

FUNCTIONAL ANALYSIS
and Other Mathematics, 2006, **1**(1), 97–103

Morse functions statistics*

Liviu I. Nicolaescu

Abstract. We prove a conjecture of V. I. Arnold concerning the growth rate of the number of Morse functions on the two-sphere.

Keywords: geometric equivalence of Morse functions; asymptotic estimates; plane graphs

MSC-2000: 05A16; 57M15

1. Introduction

We are interested in *excellent* Morse functions $f : S^2 \to \mathbb{R}$, where the attribute excellent signifies that no two critical points lie on the same level set of f. Two such Morse functions f_0, f_1 are called geometrically equivalent if there exist orientation preserving diffeomorphisms $R : S^2 \to S^2$ and $L : \mathbb{R} \to \mathbb{R}$ such that $f_1 = L \circ f_0 \circ R^{-1}$. We denote by $g(n)$ the number of equivalence classes of Morse functions with $2n+2$ critical points. Arnold suggested in [1] that

$$\lim_{n\to\infty} \frac{\log g(n)}{n \log n} = 2 . \tag{1.1}$$

The goal of this note is to establish the validity of Arnold's prediction.

Acknowledgement. I want to thank Francesca Aicardi for drawing my attention to Arnold's question.

2. Some background on the number of Morse functions

We define

$$h(n) := \frac{g(n)}{(2n+1)!} , \quad \xi(\theta) := \sum_{n\geq 0} h(n)\theta^{2n+1} .$$

In [4] we have embedded $h(n)$ in a 2-parameter family

$$(x,y) \longmapsto \hat{H}(x,y) , \quad x,y \in \mathbb{Z}_{\geq 0} , \quad h(n) = \hat{H}(0,n)$$

* This work was partially supported by NSF grant DMS-0303601.

which satisfies a nonlinear recurrence relation, [4, §8].

A. $x > 0$.

$$
\begin{aligned}
(x+2y+1)\hat{H}(x,y) &- (x+1)\hat{H}(x+1,y-1) \\
&= \frac{x+1}{2}\hat{H}(x-1,y) + \frac{x+1}{2}\sum_{(x_1,y_1)\in R_{x,y-1}} \hat{H}(x_1,y_1)\hat{H}(\bar{x}_1,\bar{y}_1)\ ,
\end{aligned}
$$

where

$$R_{x,y-1} = \{(a,b)\in\mathbb{Z}^2;\ \ 0\le a\le x,\ \ 0\le b\le y-1\}\ ,$$

and for every $(a,b)\in R_{x,y-1}$ we denoted by $(\bar{a},\bar{b})$ the symmetric of (a,b) with respect to the center of the rectangle $R_{x,y-1}$.

B. $x = 0$.

$$(2y+1)\hat{H}(0,y) - \hat{H}(1,y-1) = \frac{1}{2}\sum_{y_1=0}^{y-1}\hat{H}(0,y_1)\hat{H}(0,y-1-y_1)\ .$$

Observe that if we let $y = 0$ in **A** we deduce

$$\hat{H}(x,0) = \frac{1}{2}\hat{H}(x-1,0)\ ,$$

so that $\hat{H}(x,0) = 2^{-x}$.

In [4] we proved that these recurrence relations imply that the function

$$\xi(u,v) = \sum_{x,y\ge 0}\hat{H}(x,y)u^x v^{x+2y+1}$$

satisfies the quasi-linear PDE

$$-\left(1+u\xi+\frac{u^2}{2}\right)\partial_u\xi + \partial_v\xi = \left(\frac{1}{2}\xi^2+u\xi+1\right)\ ,\quad \xi(u,0)=0\ ,$$

and the inverse function $\xi(0,\theta) = \xi \longmapsto \theta$ is defined by the elliptic integral

$$\theta = \int_0^{\xi}\frac{\mathrm{d}t}{\sqrt{t^4/4 - t^2 + 2\xi t + 1}}\ . \tag{2.1}$$

3. Proof of the asymptotic estimate

Using the recurrence formula **B** we deduce that for every $n \geq 1$ we have

$$(2n+1)h(n) \geq \frac{1}{2}\sum_{k=0}^{n-1} h(k)h(n-1-k) \,.$$

We multiply this equality by t^{2n} and we deduce

$$\sum_{n\geq 1}(2n+1)h(n)t^{2n} \geq \frac{1}{2}\sum_{n\geq 1}\left(\sum_{k=0}^{n-1} h(k)h(n-1-k)\right)t^{2n}$$

$(g(0)=1)$

$$\Longleftrightarrow \frac{\mathrm{d}\xi}{\mathrm{d}t} \geq 1+\frac{1}{2}\xi^2.$$

This implies that the Taylor coefficients of ξ are bounded from below by the Taylor coefficients of the solution of the initial value problem

$$\frac{\mathrm{d}u}{\mathrm{d}t} = 1+\frac{1}{2}u^2\,, \quad u(0)=\xi(0)=0\,.$$

The latter initial value problem can be solved by separation of variables

$$\frac{\mathrm{d}u}{1+\frac{u^2}{2}} = \mathrm{d}t \Longrightarrow u=\sqrt{2}\tan(t/\sqrt{2})\,.$$

The function tan has the Taylor series (see [3, § 1.41])

$$\tan x = \sum_{k=1}^{\infty}\frac{2^{2k}(2^{2k}-1)|B_{2k}|}{(2k)!}x^{2k-1}\,,$$

where B_n denote the Bernoulli numbers generated by

$$\frac{t}{e^t-1} = \sum_{n=0}^{\infty} B_n\frac{t^n}{n!}\,.$$

The Bernoulli numbers have the asymptotic behavior [2, Sect. 6.2]

$$|B_{2k}| \sim \frac{2(2k)!}{(4\pi^2)^k}\,.$$

If T_k denotes the coefficient of x^{2k+1} in $\tan x$ we deduce that

$$T_k = \frac{2^{2k+2}(2^{2k+2}-1)|B_{2k+2}|}{(2k+2)!} \sim \frac{2^{2k+3}(2^{2k+2}-1)}{(4\pi^2)^{k+1}}\,.$$

Thus the coefficient u_k of t^{2k+1} in $\sqrt{2}\tan(t/\sqrt{2})$ has the asymptotic behavior

$$u_k \sim \frac{1}{2^k}\frac{2^{2k+3}(2^{2k+2}-1)}{(4\pi^2)^{k+1}} = \frac{2^{k+3}(2^{2k+2}-1)}{(4\pi^2)^{k+1}} .$$

We deduce that

$$g(k) > (2k+1)!\,\frac{2^{k+3}(2^{2k+2}-1)}{(4\pi^2)^{k+1}}(1+o(1)) \quad \text{as } k \to \infty . \tag{†}$$

Let us produce upper bounds for $g(n)$. We will give a combinatorial argument showing that

$$g(n) \leq (2n+1)!C_n ,$$

where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the n-th Catalan number.

As explained in [1, 4], a geometric equivalence class of a Morse function on S^2 with $2n+2$ critical points is completely described by a certain labelled tree, dubbed a Morse tree in [4] (see Fig. 1, where the Morse function is the height function). For the reader's convenience we recall that a Morse tree (with $2n+2$ vertices) is a tree with vertices labelled by $\{0,1,\dots\}$ and having the following two properties.

- Any vertex has either one neighbor, or exactly three neighbors, in which case the vertex is called a node.
- Every node has at least one neighbor with a higher label, and at least one neighbor with a lower label.

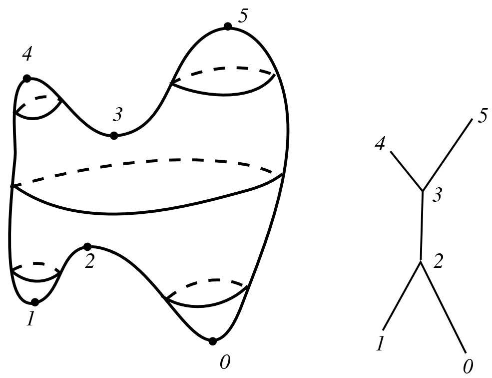

Fig. 1. Associating a tree with a Morse function on S^2

We will produce an injection from the set $\mathcal{M}_n$ of Morse functions with $2n+2$ critical points to the set $\mathcal{P}_n \times S_{2n+1}$ where $\mathcal{P}_n$ denotes the set of Planted, Trivalent, Planar Trees (PTPT) with $2n+2$ vertices, and S_{2n+1} denotes the group of permutations of $2n+1$ objects.

As explained in [4, Proposition 6.1], to a Morse tree we can canonically assign a PTPT with $2n+2$ vertices. The number of such PTPT's is C_n, [5, Exercise 6.19.f, p. 220]. The tree in Fig. 1 is already a PTPT.

The non-root vertices of such a tree can be labelled in a canonical way with labels $\{1, 2, \ldots, 2n+1\}$ (see the explanation in [5, Fig. 5.14, p. 34]). More precisely, consider a very thin tubular neighborhood N of such a tree in the plane. Its boundary is a circle. To label the vertices, walk along ∂N in the counter-clockwise direction and label the non-root vertices in the order they were first encountered (such a walk passes three times near each node). In Fig. 2, this labelling is indicated along the points marked $\circ$. The Morse function then defines another bijection from the set of non-root vertices to the same label set. In Fig. 2 this labelling is indicated along the vertices marked $\bullet$.

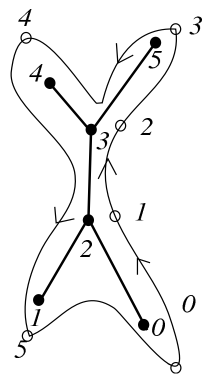

Fig. 2. Labelling the vertices of a PTPT

We have thus associated with a Morse tree a pair, (T, φ), where T is a PTPT and φ is a permutation of its non-root vertices. In Figure 2 this permutation is

$$1 \to 2\,, \quad 2 \to 3\,, \quad 3 \to 5\,, \quad 4 \to 4\,, \quad 5 \to 1\,.$$

The Morse tree is uniquely determined by this pair. We deduce that

$$\begin{aligned} g(n) &= \#\mathcal{M}_n \leq \#\mathcal{P}_n \times \#S_{2n+1} = C_n(2n+1)! \\ &= \frac{(2n)!}{(n+1)!\,n!}(2n+1)! = \frac{2 \cdot 4 \cdots (2n) \cdot 1 \cdot 3 \cdot 5 \cdots (2n-1)}{n! \cdot n!} \cdot \frac{(2n+1)!}{n+1}\,. \end{aligned}$$

Hence

$$g(n) < 2^n \cdot \frac{(2n+1)!}{n+1} \prod_{k=0}^{n-1} \frac{2k-1}{k+1} \leq \frac{2^{2n}}{(n+1)}(2n+1)!\,. \tag{‡}$$

The estimates (†) and (‡) coupled with Stirling's formula show that

$$\lim_{n\to\infty} \frac{\log g(n)}{n \log n} = 2\,,$$

which is Arnold's prediction, (1.1). □

Remark 3.1. (a) Numerical experiments suggest that

$$g(n) < (2n+1)!\,.$$

Is it possible to give a purely combinatorial proof of this inequality?

(b) It would be interesting to have a more refined asymptotic estimate for $g(n)$ of the form

$$\log g(n) = 2n\log n + r_n\,,\quad r_n = an + b\log n + c + O(n^{-1})\,,\quad a,b,c\in\mathbb{R}\,.$$

The refined Stirling's formula

$$\log(2n+1)! = \left(2n+\frac{3}{2}\right)\log(2n+1) - 2n - 1 + \frac{1}{2}\log(2\pi) + O(n^{-1})$$

implies that

$$\begin{aligned}\log h(n) &= \log g(n) - \log(2n+1)!\\ &= 2n\log n + r_n - \left(2n+\frac{3}{2}\right)\log(2n+1) + 2n + 1 - \frac{1}{2}\log(2\pi) + O(n^{-1})\\ &= r_n + 2n\left(1+\log\frac{n}{2n+1}\right) - \frac{3}{2}\log(2n+1) + 1 - \frac{1}{2}\log(2\pi) + O(n^{-1})\,.\end{aligned}$$

Hence

$$r_n = \underbrace{\log h(n) - 2n\left(1+\log\frac{n}{2n+1}\right) + \frac{3}{2}\log(2n+1) - 1 + \frac{1}{2}\log(2\pi)}_{\delta_n} + O(n^{-1})\,.$$

We deduce that

$$\frac{r_n}{n} = \frac{\delta_n}{n} + O(n^{-2})\,.$$

Here are the results of some numerical experiments.

n	δ_n/n
10	-0.634
20	-0.750
30	-0.790
40	-0.811
50	-0.824
100	-0.849
150	-0.858
200	-0.862

This suggests $a \approx -0.8\ldots$

References

1. V. I. Arnold. Smooth functions statistics. Preprint, http://www.math.jussieu.fr/seminaires/singularites/Arnold.html
2. E. A. Bender. Asymptotic methods in enumeration. *SIAM Rev.*, 1974, **16**(4), 485–515.
3. I. S. Gradshteyn, I. M. Ryzhik. Table of Integrals, Series, and Products, 6th ed. San Diego, CA: Academic Press, 2000.
4. L. I. Nicolaescu. Counting Morse functions on the 2-sphere. Preprint, math.GT/0512496.
5. R. P. Stanley. Enumerative Combinatorics, Vol. II. Cambridge Univ. Press, 1999 (Cambridge Stud. Adv. Math., 62).

Received April 24, 2006
Accepted May 15, 2006

Liviu I. Nicolaescu
Department of Mathematics
University of Notre Dame
Notre Dame, IN 46556-4618, U.S.A.
E-mail: nicolaescu.1@nd.edu
URL: http://www.nd.edu/˜lnicolae/