ANNALI DELLA
SCUOLA NORMALE SUPERIORE DI PISA
Classe di Scienze

WILHELM KLINGENBERG

Closed geodesics on surfaces of genus 0

Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4e série, tome 6, nº 1 (1979), p. 19-38

<http://www.numdam.org/item?id=ASNSP_1979_4_6_1_19_0>

Closed Geodesics on Surfaces of Genus 0 (*).

WILHELM KLINGENBERG (**)

1. – One of the outstanding results in Global Differential Geometry is the theorem of Lusternik and Schnirelmann that there exist on any closed orientable surface of genus 0 three non-selfintersecting closed geodesics. See [LS 1, 2] as well as our « Lectures on Closed Geodesics » [Kl]—briefly: LCG—for a simplified proof.

In the present paper we want to show how an extension of the methods which we employed in LCG for the proof of the Lusternik-Schnirelmann theorem yields an elementary proof for the existence of infinitely many prime (i.e., not multiply covered) closed geodesics on a surface of genus 0.

In LCG we have already presented a proof for the existence of infinitely many prime closed geodesics on an arbitrary riemannian manifold with finite fundamental group. This proof, however, is quite involved and uses rather delicate arguments. So it seems of interest to give a much simpler proof for the case of a surface of genus 0.

A fundamental rôle in our proof play the subspaces of the space of all closed curves formed by those curves having selfintersection number $\leqslant$ a fixed integer ν. We will introduce energy-decreasing deformations of those spaces into itself and show that for every $\nu = 0, 1, \ldots,$ there exist two not necessarily prime closed geodesics with selfintersection number ν. Moreover, as ν increases, the energy of these geodesics does. Cf. (5.5) for details.

2. – We begin by fixing some notations. By M we denote the 2-sphere S^2, endowed with an arbitrary riemannian metric. PM shall be the space of all

(*) This paper was written while the author was visiting the Università degli Studi and the Scuola Normale Superiore at Pisa in spring 1977 under the auspices of the Consiglio Nazionale delle Ricerche.

(**) Mathematisches Institut der Universität, 53 Bonn.

Pervenuto alla Redazione il 2 Agosto 1977.

piecewise differentiable maps (with an infinite number of geodesic pieces being allowed)

$$c\colon S = [0, 1]/\{0, 1\} \to M$$

of the parameterized circle S into M, endowed with the metric

$$d_\infty(c, c') = \sup_{t\in S} d\bigl(c(t), c'(t)\bigr) .$$

Here $d(\ ,\ )$ denotes the distance on M, derived from the riemannian metric.

On PM we have defined the energy

$$E\colon PM \to \mathbf{R}\,; \qquad c \mapsto \tfrac{1}{2}\int_S |\dot{c}(t)|^2 dt .$$

For any real $\varkappa$ we put

$$P^\varkappa M = \{c \in PM;\, E(c) \leqslant \varkappa\} .$$

In particular, $P^0 M$ consists of the constant maps $c\colon S \to M$ and thus is canonically isomorphic to M.

On PM we have the canonical S-action coming from the $S = SO(2)$-action on S:

$$(z, c) = \bigl(e^{2\pi i r}, c(t)\bigr) \in S \times PM \mapsto z\cdot c = \bigl(c(t+r)\bigr) \in PM .$$

That is to say, $z\cdot c$ is obtained from c by changing the initial point from $0 \in S$ to $r \in S$.

We also have the $\mathbf{Z}_2$-action generated by

$$\theta\colon PM \to PM\,; \qquad \bigl(c(t)\bigr) \mapsto \bigl(c(1-t)\bigr) ,$$

i.e., by the reversal of the orientation.

The S-action and the $\mathbf{Z}_2$-action together give a $O(2)$-action on PM.

The isotropy group $I^{\sim}(c)$ of $c \in PM$ under the S-action either is S (and this occurs if and only if $c \in P^0 M$), or else, is a finite cyclic group. The order of this group is called the multiplicity of c. If c has multiplicity 1 it also is called prime. If c has multiplicity m then c can be written as $c(t) = c_0(mt)$. c_0 is uniquely determined and is called the underlying prime closed geodesic of c.

We now introduce the fundamental concept of the selfintersection number $\nu(c)$ of a closed curve $c \in PM$. Assume first that $c\colon S \to M$ is an immersion and that the multiple points—if they occur at all—all are double points representing a transversal intersection. Then we define $\nu(c)$, to be the number of these double points.

For every integer $\nu \geqslant 0$ we now define $P_\nu M$ to be the closure in PM of the immersions with $\leqslant \nu$ transversal double points. Put $P_{-1}M = \emptyset$. We thus have the strictly increasing sequence

$$P_0 M \subset P_1 M \subset P_2 M \subset \ldots$$

of subspaces of PM. If $c \in P_\nu M$ but not in $P_{\nu-1} M$ then we define $\nu(c)$ to be ν.

Not every $c \in PM$ will belong to one of the $P_\nu M$. For the remaining c we put $\nu(c) = \infty$. But certainly for closed geodesics c, $\nu(c)$ is finite.

Actually, there exists the following relation between the selfintersection number of a multiply covered closed geodesic and the selfintersection number of the underlying prime closed geodesic:

2.1 PROPOSITION. *Let c be a closed geodesic of multiplicity $m(c)$. Denote by c_0 the underlying prime closed geodesic. Then*

$$\nu(c) = m(c)^2 \nu(c_0) + m(c) - 1 .$$

PROOF. We can asume: $m(c) > 1$. Observe now that the $m(c)$-fold covered closed geodesic c_0 can be approximated by curves which have $m(c)$ arcs near c_0 and $m(c) - 1$ transversal double points away from the selfintersection points of c_0. Near each of the selfintersection points of c_0 (which are transversal but may not be only double points but also points with $k > 2$ arcs passing through it) the approximating curves can be assumed to have double points only in number $m(c)^2 k$ if $k + 1$ is the number of arcs passing through that selfintersection point.

NOTE. It is the concept of the selfintersection number where it is important that M is 2-dimensional. Indeed, if dim $M > 2$, $P_0 M = PM$ and thus the selfintersection number carries no information in this case.

3. – Our next goal is to define an E-decreasing deformation of the subspace $P_\nu M$ into itself. For $\nu = 0$, Lusternik and Schnirelmann, cf. [Ly], had defined already such a deformation. This deformation seems very complicated and we had difficulties in establishing the necessary properties for it. Therefore we proposed in LCG a much simpler and actually more efficient E-decreasing deformation of $P_0 M$ into itself. Here we will show that the same deformation also can be used to transform $P_\nu M$ into itself.

As it is customary in this theory, the deformations are defined only on a subspace $P_v^\varkappa M = P_v M \cap P^\varkappa M$ of $P_v M$, $\varkappa$ some arbitrarily fixed positive real number. For such a $\varkappa$ we choose an even integer $k > 0$ such that $4\varkappa/k \leqslant \eta^2$. Here, η is a positive number such that, for every $p \in M$, the disc of radius 2η around p is a convex neighborhood of p.

Let $c \in PM$. The relation (with L denoting the length)

$$L(c|[t_0, t_1])^2 \leqslant 2E(c|[t_0, t_1])\,|t_1 - t_0|$$

implies: If $c \in P^\varkappa M$ and $|t_1 - t_0| \leqslant 2/k$ then

$$L(c|[t_0, t_1]) \leqslant \eta\,.$$

Thus in particular, there exist a unique minimizing geodesic segment $c_{c(t_0)c(t_1)}$ from $c(t_0)$ to $c(t_1)$.

We begin by defining the deformation

$$\tilde{\mathfrak{D}}_\sigma \colon P_v^\varkappa M \to P_v^\varkappa M\,; \qquad \sigma \in [0, 1/k]\,.$$

Let $c \in P_v^\varkappa M$. Put $c(0) = p$. For every $\sigma \in [0, 2/k]$ the geodesic segment $c_{pc(\sigma)}$ is well defined. The closed curve $\mathfrak{D}_\sigma c = c_{pc(\sigma)} \cup c|[\sigma, 1]$ again belongs to $P^\varkappa M$. However, it may happen that, as σ increases, $c_{pc(\sigma)}$ will have new proper (*i.e.*, transversal) intersections with $c|[\sigma, 1]$ as compared to the original arc $c|[0, \sigma]$. Whenever this begins to occur we start modifying $\sigma|[\sigma, 1]$ by substituting for each small arc of $c|[\sigma, 1]$, which comes to lie on the « wrong » side of $c'_{pc(\sigma)}$ so as to cause additional transversal selfinteractions, a geodesic segment on $c_{pc(\sigma)}$ which goes from the initial point to the end point of such an arc.

It may be necessary to make such a substitution simultaneously for several small arcs. But this is no problem, since the whole procedure takes place inside the convex neighborhood $B_{2\eta}(p)$ of p where the geodesic segments $c_{pc(\sigma)}$ look like straight segments starting from p. The substituting process may be thought of as a pushing aside or sweeping aside of parts of $c|[\sigma, 1]$ which come to lie in the way of the segment $c_{pc(\sigma)}$ as it moves around.

We denote by $\tilde{\mathfrak{D}}_\sigma c|[\sigma, 1]$ the modification of $c|[\sigma, 1]$. Since arcs of $c|[\sigma, 1]$ may be replaced by geodesic segments, the E-value of $\tilde{\mathfrak{D}}_\sigma c|[\sigma, 1]$ will be $\leqslant E(c|[\sigma, 1])$. As we proceed with σ from 0 to $2/k$ it also may happen that $c(\sigma)$ has been replaced by $\tilde{\mathfrak{D}}_\sigma c(\sigma)$. We now define

$$\tilde{\mathfrak{D}}_\sigma c = c_{p\tilde{\mathfrak{D}}_\sigma c(\sigma)} \cup \tilde{\mathfrak{D}}_\sigma c|[\sigma, 1]\,.$$

We proceed to define in the same way $\tilde{\mathfrak{D}}_\sigma c$ for $\sigma \in [2/k, 4/k]$, $c \in P^\varkappa_\nu M$: The only difference is that $c(0)$ is being replaced by $c(2/k)$ as initial point and we begin by substituting the geodesic segment $c_{c(2/k)c(\sigma)}$ for the arc $c|[2/k, \sigma]$. $\tilde{\mathfrak{D}}_{4/k}c|[2/k, 4/k]$ will be a geodesic segment of length $\leqslant \eta$.

We continue in this manner until we reach the interval $[(k-2)/k, 1]$. We then proceed to define $\tilde{\mathfrak{D}}_\sigma c$ for $\sigma \in [1, 1+2/k]$ as being a similar deformation with $c(1/k)$ as initial point and having the effect that $c|[1/k, \sigma - 1 + 1/k]$, $\sigma \in [1, 1+2/k]$ is being replaced by a geodesic segment. We go on until we reach the interval $\sigma \in [2-2/k, 2]$ where $\tilde{\mathfrak{D}}_\sigma c$ represents a curve such that $\tilde{\mathfrak{D}}_\sigma c|[1-1/k, \sigma - 1 + 1/k]$ is a geodesic segment.

We now define $\tilde{\mathfrak{D}}(\sigma, c)$, $\sigma \in [0, 2]$, to be the subsequent application of the mappings $\tilde{\mathfrak{D}}_{2/k}, \ldots, \tilde{\mathfrak{D}}_{2l/k}, \tilde{\mathfrak{D}}_\sigma$, where $2l$ is the even integer determined by $2l \leqslant k\sigma < 2l+2$.

3.1 PROPOSITION. *Choose a $\varkappa > 0$ and an integer $\nu \geqslant 0$. Then the mapping*

$$\tilde{\mathfrak{D}}: [0, 2] \times P^\varkappa_\nu M \to P^\varkappa_\nu M$$
$$(\sigma, c) \mapsto \tilde{\mathfrak{D}}(\sigma, c)$$

is continuous. Moreover, $E(\tilde{\mathfrak{D}}(2, c)) \leqslant E(c)$ with equality if and only if either c is a constant map or else a closed geodesic.

PROOF. We consider $P^\varkappa_\nu M$ to be a subspace of PM. Then the continuity is obvious. The last statement follows from the following standard facts of riemannian geometry: Let $c: [t_0, t_1] \to M$ be a piecewise differentiable map from $p = c(t_0)$ to $q = c(t_1)$, $d(p, q) \leqslant \eta$. Let c_{pq} be the unique minimizing geodesic segment from p to q, parameterized by $[t_0, t_1]$. Then $E(c_{pq}) \leqslant E(c)$ with equality if and only if $c_{pq}(t) = c(t)$.

For reference we formulate another standard result of local riemannian geometry:

3.2 PROPOSITION. *Let $\{c_n\}$ be a sequence of piecewise differentiable paths $c_n: [0, 1] \to M$. Put $c_n(0) = p_n$, $c_n(1) = q_n$. Assume: $d(p_n, q_n) \leqslant \eta$ with $\eta > 0$ as above so that there exists the uniquely determined minimizing geodesic segment $c_{p_n q_n}: [0, 1] \to M$ from p_n to q_n.*

Let now $\{E(c_n)\}$ and $\{E(c_{p_n q_n})\}$ both be convergent with the same limit. Then $\{c_n\}$ possesses a convergent subsequence with limit a minimizing geodesic segment $c: [0, 1] \to M$.

PROOF. Since M is compact, there exists a subsequence of $\{c_n\}$ which we denote again by $\{c_n\}$ such that $\lim p_n = p$ and $\lim q_n = q$ exist. It follows that $\{c_{p_n q_n}\}$ converges to the minimizing geodesic segment $c = c_{pq}$ from p to q.

Let $t_0 \in]0, 1[$. Put $c_n(t_0) = r_n$. We will show that $\lim r_n$ exists and is equal to $c(t_0)$. Indeed, $\{r_n\}$ possesses a convergent subsequence $\{r_{n(k)}\}$. Let r be its limit. The sequences

$$\{c_{p_{n(k)} r_{n(k)}}\}, \qquad \{c_{r_{n(k)} q_{n(k)}}\}$$

of unique minimizing geodesic segments converge to c_{pr} and c_{rq}, respectively, with parameter domain $[0, t_0]$ and $[t_0, 1]$. From our assumptions follows:

$$E(c_{pr}) + E(c_{rq}) \leqslant E(c_{pq}) .$$

Since $c_{pq}\colon [0, 1] \to M$ is the unique segment of minimal E-value from p to q, the curve $c_{pr} \cup c_{rq}$ from p to q must coincide with c_{pq}. In particular, its value r for $t = t_0$ must be equal to $c_{pq}(t_0)$. This completes the proof.

3.3 Lemma. *Choose $\varkappa \geqslant 0$ and an integer $\nu \geqslant 0$. Let $\widetilde{\mathfrak{D}}(\sigma, \ )$ be the deformation of $P^{\varkappa}_{\nu}M$ into itself, as defined above.*

Assume that $\{c_n\nu\}$ is a sequence in $P^{\varkappa}_{\nu}M$ such that both, $\{E(c_n)\}$ and $\{E(\widetilde{\mathfrak{D}}(2, c_n))\}$, are convergent with the same limit $\varkappa_0 > 0$. Then $\{c_n\}$ possesses a convergent subsequence with limit a closed geodesic c_0 with $E(c_0) = \varkappa_0$, $\nu(c_0) \leqslant \nu$. If, for all sufficiently large, $\nu(\widetilde{\mathfrak{D}}(2, c_n)) = \nu$ then also $\nu(c_0) = \nu$.

Proof. Besides the deformation $\widetilde{\mathfrak{D}}(\sigma, \ )$ we will also consider the deformation $\mathfrak{D}(\sigma, \ )$ of Lusternik and Schnirelmann, (cf. [Ly] and LCG) which is defined as follows: Choose k as for the definition of $\widetilde{\mathfrak{D}}(\sigma, \ )$.

For $\sigma \in [j/k, (j+2)/k]$, $j = 0, 2, \ldots, k-2$, put

$$\mathfrak{D}_\sigma c|[j/k, \sigma] = c_{c(j/k)c(\sigma)}$$
$$\mathfrak{D}_\sigma c|[\sigma, j/k+1] = c|[\sigma, j/k+1] .$$

Similarly, for $\sigma \in [1 + j/k, 1 + (j+2)/k]$, $j = 0, 2, \ldots, k-2$, define

$$\mathfrak{D}_\sigma c|[(j+1)/k, \sigma - 1 + 1/k] = c_{c((j+1)/k)c(\sigma-1+1/k)}$$
$$\mathfrak{D}_\sigma c|[\sigma - 1 + 1/k, 1 + (j+1)/k] = c|[\sigma - 1 + 1/k, 1 + (j+1)/k] .$$

Here, as always, the t-parameter has to be taken modulo 1.

Let now $\mathfrak{D}(\sigma, c)$, $\sigma \in [0, 2]$, be the subsequent application of the deformations $\mathfrak{D}_{2/k}, \ldots, \mathfrak{D}_{2l/k}, \mathfrak{D}_\sigma$, with $2l$ being the even integer determined by $2l \leqslant k\sigma < 2l + 2$.

We then have the relation

$$E(\tilde{\mathfrak{D}}(\sigma, c_n)) \leqslant E(\mathfrak{D}(\sigma, c_n)) \leqslant E(c_n) .$$

Thus we get from our hypothesis and from (3.2), since $\mathfrak{D}(1, c_n)$ consists of $k/2$ geodesic segments: There exists a subsequence of $\{c_n\}$ which we denote again by $\{c_n\}$ such that both, $\{c_n\}$ and $\{\mathfrak{D}(1, c_n)\}$ converge to a $k/2$ times broken geodesic c_0 with $E(c_0) = \varkappa_0$. Applying the same argument to the sequences

$$\{\mathfrak{D}(1, \mathfrak{D}(1, c_n))\} = \{\mathfrak{D}(2, c_n)\} \quad \text{and} \quad \{\mathfrak{D}(1, c_n)\}$$

we see: $\{\mathfrak{D}(2, c_n)\}$ converges to c_0. Since $E(\tilde{\mathfrak{D}}(2, c_0)) = E(\mathfrak{D}(2, c_0)) = E(c_0)$, c_0 is a closed geodesic, cf. (3.1).

Clearly $\nu(c_0) \leqslant \nu$, since $P_\nu^\varkappa$ is closed. If $\nu(\mathfrak{D}(2, c_n)) = \nu$, all large n, then also $\nu(c_0) = \nu$. To see this we observe that $\mathfrak{D}(2, c_n)$ consists of $k/2$ geodesic segments: $\mathfrak{D}(2, c_n)|[(j+1)/k, (j+3)/k]$ is such a segment and it converges to $c_0|[(j+1)/k, (j+3)/k]$. Hence, $\nu(\mathfrak{D}(2, c_n)) = \nu(c_0)$. But

$$\nu \geqslant \nu(c_0) = \nu(\mathfrak{D}(2, c_n)) \geqslant \nu(\tilde{\mathfrak{D}}(2, c_n)) = \nu ,$$

hence our claim. This completes the proof of (3.3).

From (3.3) we easily get

3.4 LEMMA. *Let $\varkappa_0 > 0$. Let ν be an integer $\geqslant 0$. Denote by C the set of closed geodesics c with $E(c) = \varkappa_0$, $\nu(c) \leqslant \nu$. Let $\mathfrak{U}$ be an open neighborhood of C. In the case $C = \emptyset$ one may choose $\mathfrak{U} = \emptyset$. Let $\varkappa > \varkappa_0$ and consider on $P_\nu^\varkappa M$ the deformation $\tilde{\mathfrak{D}} = \tilde{\mathfrak{D}}(2,\ )$. Then there exists a $\varepsilon > 0$ such that*

$$\tilde{\mathfrak{D}} P_\nu^{\varkappa_0+\varepsilon} \subset \mathfrak{U} \cup P_\nu^{\varkappa_0-\varepsilon} .$$

PROOF. Since $\tilde{\mathfrak{D}}$ is continuous, there exists an open neighborhood $\mathfrak{U}'$ of C, $\mathfrak{U}' \subset \mathfrak{U}$, such that $\tilde{\mathfrak{D}}\mathfrak{U}' \subset \mathfrak{U}$. If there were no $\varepsilon > 0$ with the desired property this would imply the existence of a sequence $\{c_n\}$ in $P_\nu^\varkappa M$ with $c_n \notin \mathfrak{U}'$,

$$\varkappa_0 - 1/n \leqslant E(\tilde{\mathfrak{D}} c_n) \leqslant E(c_n) \leqslant \varkappa_0 + 1/n .$$

From (3.3) we get a convergent subsequence of $\{c_n\}$ with limit a closed geodesic c_0, $E(c_0) = \varkappa_0$, $c_0 \in P_\nu$, $c_0 \notin \mathfrak{U}'$. This clearly is impossible.

4. – In this paragraph we are going to construct for every $\nu = 0, 1, \ldots,$ a rationally non-trivial cycle $w(\nu)$ of PM which actually belongs to $P_\nu M$.

In addition, we will define a certain $\boldsymbol{Z}_2$-cycle $w'(\nu)$ of the quotient space $P_\nu M/\vartheta$. In (5) this will be used to prove the existence of two closed geodesics with selfintersection number ν.

Topologically, i.e., if we disregard the distance but only consider the underlying topology, PM can be identified with PS^2. Now, the homology of PS^2 is well known. An appropriate way to describe PS^2 is to consider it as subset of the Hilbert manifold ΛS^2 of all closed H^1-curves $c\colon S \to S^2$, cf. LCG for more details. S^2 here is being endowed with the canonical metric of constant curvature equal to 1.

The energy function $\tilde{E}\colon \Lambda S^2 \to \boldsymbol{R}$ is differentiable. It satisfies the condition (C) of Palais and Smale. The set of critical points of $\tilde{E}$ on ΛS^2 decomposes into nondegenerate critical submanifolds: For $\tilde{E} = 0$, this is the manifold $\Lambda^0 S^2$ of constant maps, isomorphic to S^2. For $\tilde{E} = 2(\nu+1)^2\pi^2$, $\nu = 0, 1, 2, \ldots$, we have the manifold $B_{\nu+} S^2$ of the $(\nu+1)$-fold covered great circles. Each $B_{\nu+1} S^2$ is isometric to $B_1 S^2$, i.e., to the unit tangent bundle of S^2, which coincides with the real projective space P^3. Index $B_{\nu+1} S^2 = 2\nu + 1$.

We therefore obtain $\Lambda^{2\pi^2(\nu+1)^2} S^2$ by attaching to $\Lambda^{2\pi^2\nu^2} S^2$ the negative bundle of $B_{\nu+1} S^2$. One can show that this attaching does not affect the homology of $\Lambda^{2\pi^2\nu^2} S^2$. Since $H_i(B_{\nu+1} S^2) \neq 0$ only for $i = 0$ and 3, this attaching gives new rational cycles in dimensions $2\nu+1$ and $2\nu+4$.

For our purposes it seems preferable to describe the $(2\nu+1)$-dimensional homology class determined by the negative fibre of an element $c_{\nu+1} \in B_{\nu+1} S^2$ in a different manner: We take the standard embedding of S^2 in $\boldsymbol{R}^3 = \{(x_0, x_1, x_2)\}$. As base point $*$ of S^2 we choose $(1, 0, 0)$. Let

$$a\colon (S^2, *) \to (S^2, *)\,; \qquad b\colon (S^3, *) \to (S^2, *)$$

be the identity map and the Hopf map, respectively. To give a precise description of b we represent S^3 by

$$\{(w_0, w_1) \in \boldsymbol{C} \times \boldsymbol{C};\, w_0\overline{w}_0 + w_1\overline{w}_1 = 1\}\,.$$

The base point $*$ of S^3 shall be $(1, 0) \in \boldsymbol{C} \times \boldsymbol{C}$. Then

$$b\colon (w_0, w_1) \in S^3 \mapsto z = w_0/w_1 \in \overline{\boldsymbol{C}} = \boldsymbol{C} \cup \infty \cong S^2\,.$$

Here, the isomorphism between $\overline{\boldsymbol{C}}$ and S^2 shall be the stereographic projection of $S^2 - \{*\}$ onto $\boldsymbol{C} \cong \boldsymbol{R}^2 = \{(x, y)\}$:

$$x_0 = (x^2 + y^2 - 1)/(x^2 + y^2 + 1)\,;\; x_1 = 2x/(x^2 + y^2 + 1)\,;\; x_2 = 2y/(x^2 + y^2 + 1)\,.$$

$\infty \in \overline{\boldsymbol{C}}$ then corresponds to $* \in S^2$.

We also describe S^3 by

$$\{(e^{i\varphi}\cos\alpha,\, e^{-i\psi}\sin\alpha);\ 0\leqslant\alpha\leqslant\pi;\ 0\leqslant\varphi,\,\psi\leqslant 2\pi\}\,.$$

Then

$$b(e^{i\varphi}\cos\alpha,\, e^{-i\psi}\sin\alpha) = \big(\cos 2\alpha,\, \sin 2\alpha\cos(\varphi+\psi),\, \sin 2\alpha\sin(\varphi+\psi)\big)\,.$$

On S^3 we consider the 2-sphere S_0^2 given by $\{w_0 - \overline{w}_0 = 0\}$, i.e.

$$S_0^2 = \{(\cos\alpha,\, e^{-i\psi}\sin\alpha);\ 0\leqslant\alpha\leqslant\pi;\ 0\leqslant\psi\leqslant 2\pi\}\,.$$

Every $q\in S_0^2$ determines a parameterized circle c_q on S^3 as follows: c_q starts from $*\in S_0^2\subset S^3$ tangentially to the great circle $\{|w_0|=1\}$ into the half-sphere $\{-i(w_0-\overline{w}_0)\geqslant 0\}$ and passes for $t=\frac{1}{2}$ through the point $q\in S_0^2$.

If q is being given by $(\cos\alpha_0,\, e^{-i\psi_0}\sin\alpha_0)$ then the circle c_q is determined by the following two linear equations for u_0, v_0, u_1, v_1, with $w_0 = u_0 + iv_0$, $w_1 = u_1 + iv_1$:

$$u_1\sin\psi_0 + v_1\cos\psi_0 = 0\,; \qquad u_0\sin\alpha_0 - u_1\cos\alpha_0\cos\psi_0 + v_1\cos\alpha_0\sin\psi_0 = 0\,.$$

The image under $b\colon S^3\to S^2$ of the circle c_q therefore is the circle $\overline{b}(q)$, passing through $*\in S^2$, which in $\boldsymbol{C}$ is being represented by the straight line

$$\cos\psi_0 x + \sin\psi_0 y = \operatorname{ctg}\alpha_0\,.$$

In particular, to $\alpha_0=\pi/2$, i.e., the points q on the great circle $\{(0, e^{-i\psi})\}\subset$ $\subset S_0^2\subset S^3$ there correspond the great circles on S^2 which pass through $*\in S^2$.

We have thus defined a mapping

$$\overline{b}\colon (S_0^2, *) \to (\Omega S^2, *)$$

where ΩS^2 is the loop space of S^2. We consider only the piecewise differentiable loops starting and ending at $*\in S^2$. Thus, ΩS^2 becomes a subspace of PS^2.

We also define a map

$$\overline{a}\colon (S_0^1, *) \to (\Omega S^2, *)$$

as follows: Let S_0^1 be the great circle on S^2 given by $\{x_1 = 0\}$. Then $\overline{a}(p)$ shall be the parameterized circle starting from $*\in S_0^1\subset S^2$ into the half sphere $\{x_1\geqslant 0\}$ tangentially to the great circle $\{x_2=0\}$ and passing for $t=\frac{1}{2}$ through the point $p\in S_0^1$.

The mappings $\bar{a}$ and $\bar{b}$ are but particular explicit descriptions of the mappings $\Omega a \in \pi_1 \Omega S^2$, $\Omega b \in \pi_2 \Omega S^2$, associated to $a \in \pi_2 S^2$, $b \in \pi_3 S^2$, cf. Spanier [Sp]. Thus, $\bar{a}$, $\bar{b}$ are cycles which represent the generators of the rational homology of ΩS^2 in dimensions 1 and 2. A classical result of H. Hopf states that the rational cohomology ring of ΩS^2 is being generated by the duals $[\bar{a}]^*$, $[\bar{b}]^*$ of the homology classes $[\bar{a}]$ and $[\bar{b}]$. Moreover—and this is fundamental for our further argumentation—the products $[\bar{a}]^* \cup [\bar{b}]^{*\nu}$, $\nu = 0, 1, \ldots$, are restrictions of cohomology classes of $PS^2 \simeq \Lambda S^2$ (homotopy equivalence), cf. Švarc [Šv].

This shows that we can define cycles $w(\nu)$ of PM, dual to the $[\bar{a}^*] \cup [\bar{b}^*]^\nu$, by taking the loop product (also called: Pontrjagin product, cf. Bott-Samelson [BS]) of $\bar{a}$ with ν copies of $\bar{b}$:

$$w(\nu)\colon (p, q) \in S_0^1 \times (S_0^2)^\nu \mapsto \bar{a}(b) \cdot \bar{b}(q) \in \Omega S^2 .$$

Here, q stands for the ν-tupel $(q_1, \ldots, q_\nu)$ and $\bar{b}(q)$ denotes the loop product $\bar{b}(q_1) \cdot \ldots \cdot \bar{b}(q_\nu)$. In forming the loop product of circles we take the parameterization of the factors so as to be proportionally to arc length. I.e.:

$$L\big(w(\nu)(p, q)|[0, t_0]\big) = t_0 L\big(w(\nu)(p, q)\big) .$$

We also use $D(\nu)$ to denote the domain of $w(\nu)$. As we stated before, $w(\nu)$ is homologous to the negative fibre of the $(\nu + 1)$-fold covering of a great circle on S^2.

By γ_τ we denote the positive rotation of S_0^2 around $* \in S_0^2$ and its antipodal point $*' \in S_0^2$ by the angle τ. I.e. if q is being represented by $(\cos\alpha, e^{-i\varphi}\sin\alpha)$ then $\gamma_\tau q$ shall be given by $(\cos\alpha, e^{-i(\varphi+\tau)}\sin\alpha)$.

We denote by

$$\iota\colon S_0^2 \to S_0^2 ; \qquad (\cos\alpha, e^{-i\varphi}\sin\alpha) \mapsto (-\cos\alpha, e^{-i\varphi}\sin\alpha)$$

the reflection on the great circle $\{|w_1| = 1\} \subset S_0^2$. Then

$$\lambda = \gamma_\pi \circ \iota = \iota \circ \gamma_\pi\colon S_0^2 \to S_0^2$$

is the antipodal map.

Finally, we use γ_τ also to denote the positive rotation of $S^2 \subset \boldsymbol{R}^3$ around the points $* = (1, 0, 0)$, $*' = (-1, 0, 0)$ by the angle τ.

4.1 Proposition.

(i) $\gamma_\tau \bar{b}(q) = \bar{b}(\gamma_\tau q)$;

(ii) $\bar{b}(\lambda q) = \theta\bar{b}(q)$;

(iii) $w(\nu+1)(*, \lambda q) = \theta w(\nu+1)(*, q)$,

with $\lambda q = \lambda(q_1, \ldots, q_\nu) = (\lambda q_\nu, \ldots, \lambda q_1)$.

PROOF. (i) follows from $z(\gamma_\tau q) = e^{i\tau}z(q)$. To see (ii) we write $\bar{b}(\lambda q) = \gamma_\pi \bar{b}(\iota q)$. Let q have (α, ψ) as coordinates. The linear equation of the straight line representing the circle $\bar{b}(q)$ in the plane $\boldsymbol{C}$ is then given by

$$x \cos\psi + y \sin\psi = \operatorname{ctg}\alpha\,.$$

Therefore, $\bar{b}(\iota q)$ possesses the representation

$$x\cos\psi + y\sin\psi = -\operatorname{ctg}\alpha$$

which means that $\gamma_\pi\bar{b}(\iota q)$ coincides with $\bar{b}(q)$, up to the orientation.

To prove (iii) we use (ii):

$$w(\nu+1)(*, \lambda q) = \bar{b}(\lambda q_\nu)\cdot\ldots\cdot\bar{b}(\lambda q_1) = \\ = \theta\bar{b}(q_\nu)\cdot\ldots\cdot\theta\bar{b}(q_1) = \theta\big(\bar{b}(q_1)\cdot\ldots\cdot\bar{b}(q_\nu)\big) = \theta w(\nu+1)(*, q)\,.$$

4.2 PROPOSITION. *Let* $\tilde{E}\colon PS^2 \to \boldsymbol{R}$ *be the energy on* PS^2 *where* S^2 *has the canonical metric. Let* $\tilde{w}$ *be any cycle homological to* $w(\nu)$. *Then*

$$\sup\tilde{E}|\tilde{w} \geqslant \sup\tilde{E}|w(\nu) = 2\pi^2(2\nu+1)^2\,.$$

PROOF. Among the curves of maximal $\tilde{E}$-value in image $w(\nu)$ there is only one closed geodesic, i.e., the $(\nu+1)$-fold covering of the great circle $S_0^1 \subset S^2$. It is the image of (p_0, q_0) where $p_0 = *' = (-1, 0, 0)$ and q_0 is the ν-tupel $(0, i)^\nu$. An application of the deformation $\tilde{\mathfrak{D}}(2,\ )$ will render the $\tilde{E}$-value of all other curves in image $w(\nu) < 2\pi^2(\nu+1)^2 = \tilde{E}\big($the $(\nu+1)$-fold covering of $S_0^1\big)$.

$w(0)$ is a generator of $\pi_1\Omega S^2$. Thus, it cannot be homotopic to a $\tilde{w}$ having it image in the domain $\{\tilde{E} < 2\pi^2\}$ because then it would be homotopic to a map $S^1 \to S^2$ and thus, null-homotopic.

Therefore it only remains to prove (4.2) for $\nu > 0$. If there were a $\tilde{w} \sim w(\nu)$ with image $\tilde{w} \subset \{\tilde{E} < 2\pi^2(\nu+1)^2\}$ it would also be homotopic to a cycle representable as the negative bundle over some cycle β of the space $B_{\mu+1}S^2$ of $(\mu+1)$-fold covered great circles, $\mu < \nu$. The dimension of such a cycle is the dimension of β plus $2\mu+1$. But as observed already before, the only non-zero cycles in $B_{\mu+1}S^2$ occur in dimensions 0 and 3. $\dim\beta = 0$ is impossible, since $\mu < \nu$. And $\dim\beta = 3$ is also impossible since $(2\mu+1) + 3 \neq 2\nu+1$.

So far, the product of circles $w(\nu)(p, q) = \bar{a}(b)\cdot\bar{b}(q_1)\cdot\ldots\cdot\bar{b}(q_\nu)$ is not necessarily in $P_\nu S^2$. We show that we can replace $w(\nu)$ by a homotopic element having this property:

4.3 PROPOSITION. *The cycle* $w(\nu)$ *is homotopic* (*modulo cells of codimension* $\geqslant 2$) *to a cycle belonging to* $P_\nu S^2 \cap \{\tilde{E} \leqslant 2\pi^2(\nu+1)^2\}$. *The new cycle which we denote again by* $w(\nu)$ *still has the property* (4.1) (iii).

PROOF. As we did observe already in the proof of (4.2), the cycle $w(\nu)$ is homologous to the non-trivial cycle given by the strong unstable manifold $W_{uu}(c^{\nu+1})$ of a $(\nu+1)$-fold covered great circle $c^{\nu+1}$. This latter belongs to $P_\nu S^2 \cap \{\tilde{E} \leqslant 2\pi^2(\nu+1)^2\}$. To see this note that an eigenvector with the negative eigenvalue

$$\lambda = 4\pi^2\big(\mu^2 - (\nu+1)^2\big)/\big(1 + 4\pi^2\mu^2\big)$$

$\mu = 0, \ldots, \nu$ is given by

$$X(t) = A \quad \textit{with} \quad A \perp \dot{c}(t) \qquad \text{for } \mu = 0$$

$$X(t) = A\cos 2\pi\mu(t+\beta)\,, \qquad \text{for } 0 < \mu \leqslant \nu + 1,\ \beta \text{ real}.$$

$X(t_0) = X\big(t_0 + \varkappa/(\nu+1)\big)$, $\varkappa$ an integer, occurs at most ν times for a non-constant $X(t)$. Thus, the curves near the origin of $W_{uu}(c^{\nu+1})$ belong to $P_\nu S^2$. Along the $-\operatorname{grad} E$ flow lines the selfintersection number possibly increases. But that part of $\partial W_{uu}(c^{\nu+1})$ which goes outside $P_\nu S^2$ is homologically negligable: Indeed, from (3.3) we have that the deformation $\tilde{\mathfrak{D}}$ can be applied without encountering an obstruction since there are no critical points of selfintersection number ν below $\tilde{E} = 2\pi^2(\nu+1)$.

Consider now the cycle $w(\nu)$. The product $w(\nu)(q, p)$ of $\nu+1$ circles in general will have a selfintersection number $> \nu$. But if we apply the $-\operatorname{grad} E$ deformation we will push this cycle into the Morse complex formed by the unstable manifolds of the μ-fold covered great circles, $\mu \leqslant \nu+1$. Here, up to homologically negligable parts, $w(\nu)$ will be accomodated in $P_\nu S^2$, as we just saw. Since the $-\operatorname{grad} E$ deformation commutes with θ, the relation (4.1) (iii) is being preserved.

5. – The cycles $w(\nu)$ constructed in paragraph (4) will yield a closed geodesic $c(\nu)$ in the usual manner, using the minimum-maximum method, cf. (5.5) below. We even will find a $c(\nu)$ with selfintersection number ν. However, for the proof of the existence of infinitely many prime closed geodesics we need a second closed geodesic $c'(\nu)$ with $\nu\big(c'(\nu)\big) = \nu$.

This leads us to consider the mapping

$$w'(\nu)\colon D'(\nu) = (S_0^2)^{\nu+1} \to \Omega S^2$$
$$q = (q_0, \ldots, q_\nu) \mapsto \bar{b}(q) = \bar{b}(q_0)\cdot \ldots \cdot \bar{b}(q_\nu)\ .$$

Note: $w'(\nu) = w(\nu+1)|\{*\}\times(S_0^2)^{\nu+1}$. Moreover, by taking the modifications described in (4.3): image $w'(\nu) \subset P_\nu M$.

We define the class $W'(\nu)$ of mappings

$$w'\colon D'(\nu) \to PM$$

by the following conditions:

(i) image $w' \subset P_\nu M$;

(ii) w' is homotopic to $w'(\nu)$ by a sequence of admissible homotopies;

(iii) $w'(\lambda q) = \theta w'(q)$.

Here we mean by an admissible homotopy a continuous mapping

$$h\colon [0,1]\times D'(\nu) \to P_\nu M$$

such that, if we put $h|\{\sigma\}\times D'(\nu) = w'^{\sigma}$, $w'^{0} \in W'(\nu)$ and

$$w'^{\sigma}(\lambda q) = \theta w'^{\sigma}(q)\ .$$

5.1 Proposition. *Let $w' \in W'(\nu)$. Choose a τ_0. Then*

$$w'^{\sigma}(q) = w'^{\sigma}(q_0, \ldots, q_\nu) = w'(\gamma_{\tau_0+\pi\sigma}q_0, \ldots, \gamma_{\tau_0+\pi\sigma}q_\nu)$$

is an admissible homotopy of w'.

Proof. Immediate from $\lambda\gamma_\tau = \gamma_\tau\lambda$.

5.2 Proposition. *Let $w' \in W'(\nu)$. Then there exists a $\tilde{w}' \in W'(\nu)$, homotopic to w', such that*

$$E(\tilde{w}'(q)) \leqslant E(\tilde{\mathfrak{D}}(2, w'(q)))$$

for all $q \in D'(\nu)$. Here $\tilde{\mathfrak{D}}(2, \ )$ is a deformation from $P_\nu^\varkappa M$ into $P_\nu^\varkappa M$ of the type considered in (3). $\varkappa > 0$ has been chosen so large that image $w' \subset \{E \leqslant \varkappa\}$.

Proof. Let $H_0^2 \subset S_0^2$ be the half sphere given by $\{(\cos\alpha, e^{-i\psi}\sin\alpha);\ 0 \leqslant \alpha \leqslant \pi;\ \pi/2 \leqslant \psi \leqslant 3\pi/2\}$. Let

$$f\colon (H_0^2)^{\nu+1} \to [0,1] \subset \boldsymbol{R}$$

be a differentiable function with $f(q) = 0$ for $q \in \partial (H_0^2)^{\nu+1}$ and $f(q) = 1$ if every component of q has distance $\geqslant \varepsilon$ from ∂H_0^2, some small $\varepsilon > 0$.

Let $J(\nu + 1)$ be the set of the $2^{\nu+1}$ $(\nu + 1)$-tupels $j = (j_0, \ldots, j_\nu) \in \boldsymbol{Z}_2^{\nu+1}$. If $q = (q_0, \ldots, q_\nu) \in (S_0^2)^{\nu+1}$ we define

$$\lambda(j)\, q = (\lambda^{j_0} q_0, \ldots, \lambda^{j_\nu} q_\nu) .$$

Choose a fixed subset $J'(\nu + 1)$ of $J(\nu + 1)$ of 2^ν $(\nu + 1)$-tupels j. Then, for every $j \in J(\nu + 1)$, either $j \in J'(\nu + 1)$ or else, $j + 1 = (j_0 + 1, \ldots, j_\nu + 1) \in \in J'(\nu + 1)$.

With this we define a homotopy w'^σ of w' as follows, with $q^* \in (H_0^2)^{\nu+1}$:

$$w'^\sigma\big(\lambda(j)\, q^*\big) = \begin{cases} \tilde{\mathscr{D}}\big(2\sigma f(q^*),\, w'\big(\lambda(j) q^*\big)\big) , & \text{if } j \in J'(\nu + 1) , \\ \theta \tilde{\mathscr{D}}\big(2\sigma f(q^*), w'\big(\lambda(j + 1)\, q^*\big)\big), & \text{if } j \notin J'(\nu + 1) . \end{cases}$$

This obviously is an admissible homotopy. Moreover, the inequality (5.2) is satisfied whenever $q \in D'(\nu)$ can be written in the form $q = \lambda(j) q^*$ where $q^* \in (H_0^2)^{\nu+1}$ and $f(q^*) = 1$.

To obtain (5.2) also for those $q = \lambda(j) q^*$ where $f(q^*) < 1$ we apply to $(H_0^2)^{\nu+1}$ first one of the $(\nu + 2)$ rotations

$$\gamma_{\tau(\mu)} , \qquad \tau(\mu) = \pi\mu/(\nu + 2) ; \qquad \mu = 0, \ldots, \nu + 1$$

and follow this up with the previously defined homotopy. For every $q^{**} \in (H_0^2)^{\nu+1}$ it will happen at least once that, if $\gamma_{\tau(\mu)} q^{**} = \lambda(j) q^*$, $q^* \in \in (H_0^2)^{\nu+1}$, then $f(q^*) = 1$. This completes the proof of (5.2).

From our description of the Hopf map $b\colon S^3 \to S^2$ in (4) we can see: If $H_0^1 \subset S_0^2$ denotes the half great circle

$$H_0^1 = \{(\cos\alpha, -i \sin\alpha);\ 0 \leqslant \alpha \leqslant \pi\}$$

then $b|H_0^1\colon H_0^1 \to S^2$ has as image the great circle $S_0^1 \subset S^2$. Actually, $b|H_0^1$ is injective, with the exception of the boundary points $* \in S^3$ and $*' \in S^3$ of H_0^1 which both are mapped into $* \in S_0^1 \subset S^2$.

Therefore we may write the cycle $\bar{a}\colon S_0^1 \to \Omega S^2$ also in the form $\bar{b}|H_0^1\colon H_0^1 \to \Omega S^2$. Consequently we denote also $H_0^1 \times (S_0^2)^\nu$ by $D(\nu)$. Then

$$w(\nu) = w'(\nu)|D(\nu) .$$

This leads us to define the class $W(\nu)$ as being mappings

$$w\colon D(\nu) \to PM$$

where $w = w'|D(\nu)$, $w' \in W'(\nu)$. Clearly, any $w \in W(\nu)$ is homotopic to $w(\nu)$.

Let $w \in W(\nu)$. We call $(p, q) \in D(\nu)$ a regular point if it is in the domain of a non-degenerate singular simplex of the map $w: D(\nu) \to PM$, considered as sum of singular simplices. Similarly we define a regular point of $w' \in W'(\nu)$.

5.3 PROPOSITION. *The set of regular points of $w \in W(\nu)$ belonging to $w^{-1}(P_\nu M - P_{\nu-1} M)$ is non-empty.*

PROOF. We can assume: $\nu > 0$. Since w is a non-trivial cycle there are regular points. If $w^{-1}(P_\nu M - P_{\nu-1} M)$ were empty, i.e., if image $w \subset P_{\nu-1} M = P_{\nu-1} S^2$, then image w could be deformed in $P_{\nu-1} S^2$ below the $\tilde{E}$-level $2\pi^2(2\nu+1)^2$, since there are no closed geodesics in $P_{\nu-1} S^2$ with $\tilde{E}$-value $> 2\pi^2(2\nu-1)^2$. But this contradicts (4.2). The degenerate simplices in w are negligable homologically. Therefore, since w is non-trivial, the same argument shows that there are non-degenerate simplices in $P_\nu M - P_{\nu-1} M$.

5.4 PROPOSITION. *The set of regular points of $w' \in W'(\nu)$ having image in $P_\nu M - P_{\nu-1} M$ is non-empty.*

PROOF. Since $w' \in W'(\nu)$ is not a non-trivial rational cycle we cannot employ the same arguments as for the proof of (5.3). Instead, we will use the fact that a $w' \in W'(\nu)$ determines a non-trivial $\boldsymbol{Z}_2$-cycle of the space PM/θ of unoriented parameterized closed curves on M. Here, PM/θ is the quotient of PM with respect to the $\boldsymbol{Z}_2$-action generated by the orientation reversing map θ. Note: $\theta P_\nu M = P_\nu M$.

To make this more precise we consider the quotient space $D'(\nu)/\lambda$ of $D'(\nu)$ with respect to the $\boldsymbol{Z}_2$-action generated by $\lambda: D'(\nu) \to D'(\nu)$, cf. (4). Using the notations employed in the proof of (5.2) we can represent the fundamental $\boldsymbol{Z}_2$-homology class of the space $D'(\nu)/\lambda$ by the cycle

$$z_{2\nu+2}: Z^{2\nu+2} = \bigcup_{j \in J'(\nu+1)} \lambda(j)(H_0^2)^{\nu+1} \xrightarrow[/\lambda]{} D'(\nu)/\lambda .$$

The 1-dimensional $\boldsymbol{Z}_2$-homology class is being represented by the cycle

$$z_1: Z^1 = H_0^1 \xrightarrow[/\lambda]{} D'(\nu)/\lambda .$$

Here we have considered H_0^1 as subset of the first factor of $D'(\nu) = (S_0^2)^{\nu+1}$. Finally, the 1-codimensional $\boldsymbol{Z}_2$-homology class of $D'(\nu)/\lambda$ is represented by the cycle

$$z_{2\nu+1}: Z_{2\nu+1} = H_0^1 \times Z^{2\nu} \xrightarrow[/\lambda]{} D'(\nu)/\lambda$$

with $H_0^1 \times Z^{2\nu} \subset S_0^2 \times (S_0^2)^\nu = D'(\nu)$.

$D'(\nu)/\lambda$ has the important property that any 1-cycle and any $(2\nu+1)$-cycle, when in general position, intersect in an odd number of points. That is to say, $z_{2\nu+1}$ represents the 1-dimensional $\boldsymbol{Z}_2$-cohomology class of $D'(\nu)/\lambda$. Note that $(z_1, z_{2\nu+1})$ are not in general position.

An element $w' \in W'(\nu)$ can be viewed as the 2-fold covering of a mapping

$$w'/\lambda\colon D'(\nu)/\lambda \to P_\nu M/\theta .$$

Assume now that w' has no regular points. Since $w'|D(\nu) \in W(\nu)$ is a non-trivial cycle this implies that some 1-cycle $z_1' \sim z_1$ becomes a trivial cycle under $w'|z_1'$. But this is impossible since $w'|z_1'$ is homologous to $w(0)$ which is a non-trivial cycle, representing the generator of $H_1 \Omega S^2 = H_1 P S^2$, cf. (3).

That the image of the interior of the set of regular points meets $P_\nu M - P_{\nu-1} M$ is proved now just as in (5.3).

We now can define:

$$\varkappa(\nu) = \inf_{w \in W(\nu)} \sup \{E(c); \nu(c) = \nu, c = w(p, q), (p, q) \text{ regular point}\};$$

$$\varkappa'(\nu) = \inf_{w' \in W'(\nu)} \sup \{E(c); \nu(c) = \nu, c = w'(q), q \text{ regular point}\} .$$

5.5 THEOREM. *For every $\nu = 0, 1, \ldots,$ there exist closed geodesics $\{c(\nu), c'(\nu)\}$ with selfintersection number ν and $E(c(\nu)) = \varkappa(\nu)$, $E(c'(\nu)) = \varkappa'(\nu)$. Moreover*

$$(*) \qquad 0 < \varkappa(\nu) \leqslant \varkappa'(\nu) < \varkappa(\nu+1) .$$

Here equality can hold only, if there are infinitely unparameterized closed geodesics with E-value $\varkappa(\nu) = \varkappa'(\nu)$ and selfintersection number ν.

PROOF. Since $w \in W(\nu)$ contains as restriction $w|H_0^1 \times \{* \ldots *\}$ an element of $W(0)$, $\varkappa(\nu) > 0$ will follow if we show: $\varkappa(0) > 0$. But this can easily be seen as follows: If $\varkappa(0) = 0$ then there would exist a $w \in W(0)$ with image $w \subset \{E \leqslant \eta^2/2\}$, $\eta > 0$ as in (3). For every $p \in H_0^1$, the closed curve $w(p)$ therefore can be retracted along the radii of the convex 2η-neighborhood around $w(p)(0) = w(p)(1)$ into $w(p)(0)$. Thus, $w(0)$, being homotopic to w, becomes homotopic to a map $S^1 \to M = S^2$, i.e., to a constant map, which is a contradiction.

Let $q \in D'(\nu)$ be a regular point of a $w' \in W'(\nu)$: Then there exists a $w \in W(\nu)$ containing q among its domain of definition. Thus, q is regular point also for some $w \in W(\nu)$. This shows: $\varkappa(\nu) \leqslant \varkappa'(\nu)$.

To prove the existence of a closed geodesic $c(\nu)$ with $E(c(\nu)) = \varkappa(\nu)$,

$\nu(c(\nu)) = \nu$ we observe that we get from the definition of $\varkappa(\nu)$ a sequence $\{w_n\}$ in $W(\nu)$ and a sequence $\{r_n\}$ in $D(\nu)$, r_n interior regular point of w_n, such that for $c_n = w_n(r_n)\colon \nu(c_n) = \nu$ and

$$\varkappa(\nu) \leqslant E(\widetilde{\mathfrak{D}}(2, c_n)) \leqslant E(c_n)$$

and $\lim E(c_n) = \varkappa(\nu)$. According to (3.3) we can assume that $\lim c_n$ exists and is a closed geodesic $c(\nu)$. We can assume that also $\widetilde{\mathfrak{D}}(2, c_n)$ is in the image of an interior regular point of $\widetilde{\mathfrak{D}}(2, w_n) \in W(\nu)$ and belongs to $P_\nu M - P_{\nu-1} M$. (3.3) then implies: $\nu(c(\nu)) = \nu$.

The same arguments lead to the existence of a $c'(\nu)$ with the desired properties.

To prove $\varkappa'(\nu) < \varkappa(\nu+1)$ we first show that $\varkappa'(\nu) \leqslant \varkappa(\nu+1)$. To see this we deduce from our previous arguments the existence of a sequence $\{w_n\} \in$ $\in W(\nu+1)$, hence, $\{w_n' := w_n|\{*\} \times D'(\nu)\} \in W'(\nu)$, and of a $q \in D'(\nu)$ being regular for all w_n' such that the sequence $\{c_n = w_n'(q) = w_n(*, q)\}$ converges to $c'(\nu)$ with $\nu(c_n) = \nu$.

We claim that for any $q \in D'(\nu)$, regular for w_n', there exist $p \in H_0^1$, arbitrarily close to $* \in H_0^1$ such that $(p, q) \in D(\nu+1)$ is regular for all w_n. This means that $(*, q) \in D(\nu+1)$ is a boundary regular point for w_n. Indeed, the cycle $w_n|H_0^1 \times \{q\}$ is homotopic to $w_n|H_0^1 \times \{*\} \in W(0)$ and hence nontrivial. Therefore, after possibly some homotopic modification of w_n near $(*, q)$, $w_n|H_0^1 \times \{q\}$ will be regular at $(*, q)$, thus our claim.

Consider now $w_n|\{p\} \times D'(\nu)$. As long as the image is in $P_\nu M$, while $p \in H_0^1$ moves away from $*$, we get an element of $W'(\nu)$ and thus, $\sup E|(w_n|\{p\} \times$ $\times D'(\nu)) \geqslant \varkappa'(\nu)$. But it is impossible that image $w_n|\{p\} \times D'(\nu) \subset P_\nu M$, for all $p \in H_0^1$, cf. (5.3). Hence, there is one $p_0 \in H_0^1$ (possibly $p_0 = *$) such that p_0 is the limit of points $p \in H_0^1$ with image $w_n|\{p\} \times D'(\nu) \cap P_{\nu+1} M \neq \emptyset$. Thus, $\varkappa(\nu+1) \geqslant \varkappa'(\nu)$.

To see that actually $\varkappa'(\nu) < \varkappa(\nu+1)$ we note that the set $C(\nu)$ of closed geodesics c with $E(c) = \varkappa'(\nu)$ and $\nu(c) = \nu$ and the set $C(\nu+1)$ of closed geodesics c with $E(c) = \varkappa'(\nu)$ and $\nu(c) = \nu+1$, if not empty, are disjoint compact sets. Since we just showed that in every neighborhood of $c(\nu) \in$ $\in C(\nu)$ there are elements $c_n = w_n(p, q) \in P_{\nu+1} M - P_\nu M$, (p, q) regular for $w_n \in W(\nu+1)$, we see that $\varkappa(\nu) = \varkappa(\nu+1)$ is impossible.

It remains to discuss the case $\varkappa(\nu) = \varkappa'(\nu) =$ (briefly) $\varkappa_0$. We want to derive a contradiction from the assumption that the set C of closed geodesics c with $E(c) = \varkappa_0$, $\nu(c) = \nu$ consists of only finitely many S-orbits $S.c$. In this case we could choose arbitrarily small open neighborhoods $\mathfrak{U}$ of C of the following type: $\mathfrak{U}$ is the union of finitely many pairwise disjoint S-invariant open neighborhoods $\mathfrak{U}(S.c)$ of the finitely many S-orbits $S.c$. We also may assume:

$$\mathcal{U}(S.\theta c) = \theta\mathcal{U}(S.c)\,.$$

From (3.4) and (5.2) we have the existence of an $\varepsilon > 0$ and a $w' \in W'(\nu)$ such that

$$\text{image } w' \subset \mathcal{U} \cup P_\nu^{\varkappa_0-\varepsilon} M\,.$$

$\varkappa(\nu) = \varkappa'(\nu)$ means that the image under w' of every cycle $z'_{2\nu+1} \sim z_{2\nu+1}$ (cf. the proof of (5.4)) meets $\mathcal{U}/\theta$. Or equivalently: The carrier of every non-trivial 1-dimensional $\boldsymbol{Z}_2$-cocycle meets $\mathcal{U}/\theta$. That is to say, there is a $z'_1 \sim z_1$ where $w'(z'_1) \subset \mathcal{U}(S.c)/\theta$, some S-orbit $S.c$ in C. But $S.c$ can be retracted into a trivial S-orbit $\subset P^0 M = M$, thus, the cycle $w'|z'_1$ is trivial which is a contradiction since it is homologous to the cycle $w(0)$.

This completes the proof of (5.5).

6. – We can now prove our main result.

6.1 THEOREM. *Let M be a closed surface of genus 0. Then there exist on M infinitely many unparameterized prime closed geodesics.*

NOTES. 1) Lusternik and Schnirelmann l.c. had proved the existence of three closed geodesics without selfintersections on an orientable surface of genus 0. The ellipsoid with three pairwise different axes, all approximately of the same length, gives an example of such a surface where there exist exactly three closed geodesics without selfintersections, i.e., the three principal ellipses of the ellipsoid.

As was shown by Morse [Mo], cf. also LCG, the next prime closed geodesic in E-value and hence in length, after these three relatively short closed geodesics, can have an E-value greater than any prescribed number, if only the three axes have their length sufficiently near 1. This fourth prime closed geodesic therefore also will have arbitrarily large selfintersection number.

2) The theorem is a special case of our theorem (4.3.5) in LCG which states that on every compact riemannian manifold with finite fundamental group there exist infinitely many prime closed geodesics. The proof of this more general theorem uses various deep results, among others the Gromoll-Meyer theorem [GM], Sullivan's theory of the minimal model [Su] and the structure of the Morse complex of the Hilbert manifold of closed H^1-curves with respect to a non-degenerate energy function E, cf. LCG for details.

In contrast, the present proof of this result for a closed surface of genus 0 can be considered elementary. We could have shortened it were it not for an attempt to keep the topological prerequisites to a minimum.

PROOF. We can assume M is orientable. That is M is the 2-sphere with an arbitrary riemannian metric. Indeed, otherwise the universal covering $\tilde{M}$ of M is of this type. If we have infinitely many prime closed geodesics on $\tilde{M}$ then their projection into M under the covering map $\tilde{M} \to M$ will yield infinitely many prime closed geodesics on M also since the image of different prime closed geodesics has different underlying prime closed geodesics.

We now will derive a contradiction from the assumption that there are on M only finitely many S-orbits of prime closed geodesics. Let $\{c_1, \dots, c_s\}$ representatives of these different S-orbits.

From (5.5) we get for every $\nu = 0, 1, \dots$ a pair $\{c(\nu), c'(\nu)\}$ of closed geodesics with selfintersection number ν. Under our assumption the relations $(*)$ in (5.5) become

$$0 < E(c(\nu)) < E(c'(\nu)) < E(c(\nu + 1)) .$$

For a fixed ν there exist integers $a(\nu)$, $b(\nu)$, $0 \leqslant a(\nu) < b(\nu) \leqslant s$, such that both, $c(\nu + a(\nu))$ and $c(\nu + b(\nu))$ have the same underlying prime closed geodesic. Here and in the following we identify geodesics which lie on the same S-orbit.

Taking iterated subsequences we see: There exists a strictly increasing sequence $\{\nu(k)\}$ of integers $\geqslant 0$ and a, b, $0 < b \leqslant s$, such that, for all $k \in N$, $c(\nu(k))$ and $c(\nu(k) + b)$ have the same underlying prime closed geodesic, say c. Moreover, also the geodesics $c'(\nu(k))$ all have the same underlying prime closed geodesic, say c'.

Put $E(c) = \varkappa$ $E(c') = \varkappa'$. Denote the multiplicity of $c(\nu(k))$, $c'(\nu(k))$ and $c(\nu(k) + b)$ by $m(k)$, $m'(k)$, and $\tilde{m}(k)$, respectively. Put $\nu(c) = \nu$, $\nu(c') = \nu'$. Then we have the following relations, cf. (5.5), (2.1):

$$(*) \qquad m(k)^2 \varkappa < m'(k)^2 \varkappa' < \tilde{m}(k)^2 \varkappa$$

$$(**) \qquad \begin{cases} \text{(i)} \;\; \nu(k) + 1 = m(k)^2 \nu + m(k) = m'(k)^2 \nu' + m'(k) \\ \text{(ii)} \;\; \nu(k) + b + 1 = \tilde{m}(k)^2 \nu + \tilde{m}(k) . \end{cases}$$

This shows that with $k \to \infty$ also $m(k)$, $m'(k)$, $\tilde{m}(k) \to \infty$. Moreover:

$$(***) \qquad \lim_k m(k)^2/\tilde{m}(k)^2 = 1 \; ; \qquad \lim_k m'(k)^2/m(k)^2 = \varkappa/\varkappa' .$$

This follows from $(*)$ together with $(**)$.

If $\nu = \nu' = 0$, $m(k) = m'(k)$, $\varkappa = \varkappa'$, which contradicts $(*)$.

If $\nu > 0$ we have from

$$1 = \frac{m(k)^2(\nu + 1/m(k))}{m'(k)^2(\nu + 1/m'(k))}$$

and $(***)$:

$$(****) \qquad \varkappa/\varkappa' = \nu/\nu'.$$

From $(*)$ and $(**)$ follows:

$$m(k)^2\nu < m'(k)^2\nu' < \tilde{m}(k)^2\nu$$

$$\tilde{m}(k) - b < m'(k) < m(k) .$$

Thus, with $(***)$, $\varkappa = \varkappa'$, $\nu = \nu'$, which again contradicts $(*)$. This completes the proof.

REFERENCES

[BS] R. Bott - H. Samelson, *On the Pontryagin product in spaces of paths*, Comm. Math. Helv., **27** (1953), pp. 320-337.

[GM] D. Gromoll - W. Meyer, *Periodic geodesics on compact Riemannian manifolds*, J. Diff. Geom., **3** (1969), pp. 493-510.

[Kl] W. Klingenberg, *Lectures on Closed Geodesics*, Gruadlehren Math. Wiss. Bd. 230 Springer, Berlin-Heidelberg-New York, 1968.

[LS 1] L. Lusternik - L. Schnirelmann, *Sur le problème de trois géodésiques fermées sur les surfaces le genre* 0, C. R. Acad. Sci. Paris, **189** (1929), pp. 269-271.

[LS 2] L. Lusternik - L. Schnirelmann, *Topological methods in the calculus of variations*, Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow, 1930 (Russian); *Méthodes topologiques dans les problèmes variationelles.*, 1-re partie, Actualités Sci. Industr., no. 188, Paris, 1934.

[Ly] L. Lyusternik, *The topology of function spaces and the calculus of variations in the large*, Trudy. Mat. Inst. Steklov, **19** (1947) (Russian); *The Topology of the Calculus of Variations in thel arge*, Translations of Mathematical Monographs, vol. **16**, Amer. Math. Soc., Providence, R.I., 1966.

[Mo] M. Morse, *The calculus of variations in the large*, Amer. Math. Soc. Colloq. Publ., vol. **18**, Amer. Math. Soc., Providence, R.I., 1934.

[Sp] E. Spanier, *Algebraic topology*, McGraw Hill, New York, 1966.

[Sv] A. S. Švarc, *Homology of the space of closed curves*. Trudy Moskov, Mat. Obšč., **9** (1960), pp. 3-44 (Russian).

[Su] D.Sullivan, *Differential forms and the topology of manifolds*, Manifolds-Tokyo, 1973, ed. A. Hattori, Univ. of Tokyo Press, Tokyo, 1975.