EXPLOSIONS NEAR ISOLATED UNSTABLE ATTRACTORS

Konstantin Athanassopoulos

We describe the set of explosive orbits in the region of attraction of an unstable attractor which is isolated in the sense of C.C. Conley. Sufficient conditions are given for the existence of explosions in certain parts of the region of attraction and for an unstable attractor to have finitely generated integral Alexander-Spanier cohomology groups. Finally, we study the case of singularities that are unstable attractors in flows on the 2-sphere.

1. Introduction.

One of the most studied parts of the phase space of a continuous flow on a separable, locally compact, metrizable space $M$ is the region of attraction $W$ of an asymptotically stable compact invariant set $A$, that is a Lyapunov stable attractor. It is well-known (see [2, Section 10], [3, Theorem V.2.9], [8]) that there exists a strictly decreasing along the orbits in $W\setminus A$ uniformly unbounded Lyapunov function $f : W \to \mathbb{R}^+$ with $f^{-1}(0) = A$. For any $c > 0$, the level set $f^{-1}(c)$ is a compact global section to the restricted flow in $W\setminus A$, which is therefore parallelizable. If $M$ is a locally compact ANR, $A$ has the shape of the compact ANR $f^{-1}([0, c])$, which has the homotopy type of a compact polyhedron, and which is a strong deformation retract of $W$. Thus, $A$ and $W$ have isomorphic and finitely generated Alexander-Spanier cohomology (see [9]).

On the other hand, not much is known about the structure of the flow in the region of attraction of an unstable attractor. In this case there might be orbits in $W \setminus A$ whose positive prolongational limit set is not contained in $A$, and so the flow explodes at these orbits. In general we say that the flow explodes at an orbit if the positive limit set of the orbit is not equal to its positive prolongational limit set.

This work has two parts. In the first part we study the explosions of the flow in the region of attraction of an unstable attractor, under the additional assumption that it is an isolated invariant set in the sense of C.C. Conley. First we show in Theorem 3.4 that if $A$ is an invariant continuum and an isolated unstable attractor such that $W \setminus A$ is connected, then there exist explosions in $W \setminus D^+(A)$ or $D^+(A) = M$, where $D^+(A)$ is the positive first
prolongation of $A$. The set of orbits in $W \setminus A$ at which the flow explodes is completely determined by a suitable compact subset of the boundary of an isolating block of $A$.

The positive first prolongation $D^+(A)$ of an unstable attractor $A$ is an asymptotically stable compact invariant set with the same region of attraction as $A$ (see [2, Theorem 8.20]). Simple examples show that the flow may or may not explode at some orbit in $D^+(A) \setminus A$. In Theorem 3.5 we give a sufficient topological condition on the region of attraction in order the flow to explode at some orbit in $D^+(A) \setminus A$. We derive also some topological properties of a connected isolated unstable attractor in case there is essentially no explosion in $D^+(A) \setminus A$ and $D^+(A) \setminus A$ is topologically not too bad. More precisely, we prove in Theorem 3.7 that if $D^+(A) \setminus A$ is an ANR and the first positive prolongational limit set of every point in $D^+(A) \setminus A$ is contained in $A$ then the integral Alexander-Spanier cohomology groups of $A$ are finitely generated. The assumption that $D^+(A) \setminus A$ is an ANR is of course restrictive in general.

In the second part we study the case where $A$ is a fixed point of a flow on a 2-manifold $M$, which is an isolated unstable attractor. In this case the set of orbits in $W \setminus A$ at which the flow explodes is finite, as we prove in Corollary 4.3. If $M$ is the 2-sphere, then we prove in Theorem 4.5 that $D^+(A) \setminus A$ has finitely many connected components and each one of them is either an orbit homeomorphic to $\mathbb{R}$ or a noncompact 2-manifold with nonempty and noncompact boundary, and with finitely many boundary components. Hence $D^+(A) \setminus A$ is an ANR. Moreover, as we prove in Theorem 4.7 the restricted flow in $D^+(A) \setminus A$ is completely unstable. These are not local results, as global properties of the flow have to be taken into account.

2. Explosions and isolated invariant sets.

Let $\phi$ be a continuous flow on a separable, locally compact, metrizable space $M$. We shall denote by $\phi(t,x) = tx$ the translation of the point $x \in M$ along its orbit in time $t \in \mathbb{R}$. We shall also write $\phi(I \times A) = IA$, for $I \subset \mathbb{R}$ and $A \subset M$. The orbit of $x$ will be denoted by $C(x)$, its positive semi-orbit by $C^+(x)$ and the negative by $C^-(x)$. The positive limit set of $x \in M$ is the closed, invariant set

$$L^+(x) = \{ y \in M : t_n x \rightarrow y \text{ for some } t_n \rightarrow +\infty \}$$

and describes the behavior at infinity of $C^+(x)$. The behavior of the flow at infinity near the point $x \in M$ is described through its (first) positive prolongational limit set

$$J^+(x) = \{ y \in M : t_n x_n \rightarrow y \text{ for some } x_n \rightarrow x \text{ and } t_n \rightarrow +\infty \}$$

which is also closed and invariant. The positive first prolongation of $x$ is the positively invariant closed set $D^+(x) = C^+(x) \cup J^+(x)$. The negative
versions are defined by reversing time. Obviously, $L^+(x) \subset J^+(x)$. We say that the flow \textit{explodes} at the orbit of the point $x \in M$ in positive time if $J^+(x) \neq L^+(x)$.

A compact invariant set $A \subset M$ is called \textit{isolated} if it has a compact neighbourhood $V$ such that $A$ is the maximal invariant set in $V$. Each such $V$ is called an isolating neighbourhood of $A$. It is known (see [4], [5]) that every isolating neighbourhood of $A$ contains a smaller isolating neighbourhood $N$ of $A$ such that there are compact sets $N^+, N^- \subset \partial N$ with the following properties:

(i) $\partial N = N^+ \cup N^-$. 
(ii) For every $x \in N^+$ there exists $\epsilon > 0$ such that $[-\epsilon, 0)x \subset M \setminus N$ and for every $y \in N^-$ there exists $\delta > 0$ such that $(0, \delta)y \subset M \setminus N$.
(iii) For every $x \in \partial N \setminus N^+$ there exists $\epsilon > 0$ such that $[-\epsilon, 0)x \subset \text{int} N$ and for every $y \in \partial N \setminus N^-$ there exists $\delta > 0$ such that $(0, \delta)y \subset \text{int} N$.

The triad $(N, N^+, N^-)$ is called an \textit{isolating block} of $A$. The set $N^+$ is the \textit{entrance set} and $N^-$ is the \textit{exit set} of the isolating block. The sets $A^\pm = \{x \in N : C^\pm(x) \subset N\}$ and $\alpha^\pm = \partial N \cap A^\pm$ are compact and $A = A^+ \cap A^-$. Moreover, $\emptyset \neq L^+(x) \subset A$ for every $x \in A^+$ and $\alpha^+ \subset N^+ \setminus N^-$.

If $M$ is a smooth $n$-manifold and the flow is smooth, then every neighbourhood of an isolated invariant set $A$ contains a smooth isolating block $(N, N^+, N^-)$ of $A$. This means that $N$ is a smooth compact $n$-manifold with boundary $\partial N = N^+ \cup N^-$, the sets $N^+$ and $N^-$ are smooth compact $(n-1)$-manifolds with common boundary $N^+ \cap N^-$, which is a smooth compact $(n-2)$-manifold (without boundary) and on which the flow is externally tangent to $N$. Moreover, the flow is transverse to $N^+ \setminus N^-$ into $N$ and transverse to $N^- \setminus N^+$ out of $N$ (see [6]).

\textbf{Lemma 2.1.} Let $A$ be an isolated invariant set in $M$ and $(N, N^+, N^-)$ be an isolating block of $A$.

(a) If $x \in \text{int}_{\partial N} \alpha^+$, then $\emptyset \neq J^+(x) \subset A$.
(b) If $x \in \partial_{\partial N} \alpha^+$, then $J^+(x) \cap \alpha^- \neq \emptyset$.

\textbf{Proof.} (a) Let $W$ be an open neighbourhood of $x$ in $M$ such that $W \cap \partial N \subset \alpha^+$. Let $y \in J^+(x)$ and suppose that $y \notin A$. There are open sets $U$ and $Y$ in $M$ such that $A \subset U \subset \text{int} N$, $y \in Y$ and $\overline{U} \cap \overline{Y} = \emptyset$. Since $\emptyset \neq L^+(x) \subset A$, there exists $t > 0$ such that $C^+(tx) \subset U$. There are $t_n > t$, $x_n \in W \cap \partial N$, $n \in \mathbb{N}$, such that $t_n \to +\infty$, $x_n \to x$ and $t_n x_n \to y$. So, there are $s_n > t_n$ such that $s_n x_n \in \partial U$, $n \in \mathbb{N}$. By compactness, the sequence $\{s_n x_n : n \in \mathbb{N}\}$ has a limit point $z \in \partial U$. For every $s \in \mathbb{R}$ we have eventually $(s + s_n)x_n \in C^+(x_n)$. Therefore, $sz \in N$ for every $s \in \mathbb{R}$. This contradicts our hypothesis that $A$ is isolated.

(b) There is a sequence $\{x_n : n \in \mathbb{N}\}$ of points in $N^+ \setminus \alpha^+$ converging to $x$. Let $\sigma^+(x_n)$ be the time $C^+(x_n)$ exits $N$. Then $\sigma^+(x_n)x_n \in N^- \setminus \alpha^-$ and
\( \sigma^+(x_n) \rightarrow +\infty \), because \( x \in \alpha^+ \). By compactness, the sequence \{\( \sigma^+(x_n)x_n : n \in \mathbb{N} \)\} has a limit point \( y \in J^-(x) \cap N^- \). For each \( s < 0 \) we have eventually \( 0 < \sigma^+(x_n) + s < \sigma^+(x_n) \) and therefore \( (\sigma^+(x_n) + s)x_n \in N \). It follows that \( sy \in N \) for every \( s < 0 \), that is \( y \in J^+(x) \cap \alpha^- \).

**Remark 2.2.** Let \( A \) be an isolated invariant set and let \( x \in M \). Let \((N_i, N_i^+, N_i^-), i = 1, 2, \) be two isolating blocks of \( A \) such that \( N_1 \subset N_2 \).

Then, \( J^+(x) \cap \alpha^-_2 \) is homeomorphic to \( J^+(x) \cap \alpha^-_1 \). Indeed, for every \( x \in \alpha^-_2 \) there is a unique \( \tau(x) \leq 0 \) such that \( \tau(x)x \in \alpha^-_1 \). It is easy to see that the function \( \tau : \alpha^-_2 \rightarrow \mathbb{R} \) is continuous. Define now the continuous injection \( h : \alpha^-_2 \rightarrow \alpha^-_1 \) by \( h(x) = \tau(x)x \). This is also onto, because for every \( z \in \alpha^-_1 \) we have \( C^+(z) \notin N_2 \). Thus \( h \) is a homeomorphism, by compactness. It is now obvious that \( h(J^+(x) \cap \alpha^-_2) = J^+(x) \cap \alpha^-_1 \). If we have two isolating blocks of the isolated invariant set \( A \), then we can always find a smaller isolating block. So, the homeomorphism type of \( J^+(x) \cap \alpha^- \) does not depend on the choice of the isolating block. If \( x \in M \) is such that \( \emptyset \neq L^+(x) \subset A \), it is contained in \( J^+(x) \setminus L^+(x) \) and its size and topological complexity may give information on the explosion of the flow at \( x \). However, it seems to be difficult to deal with the set \( J^+(x) \cap \alpha^- \).

### 3. Isolated unstable attractors.

Let again \( \phi \) be a continuous flow on a separable, locally compact, metrizable space \( M \) and \( A \subset M \) be a compact invariant set. The invariant set

\[
W^+(A) = \{ x \in M : \emptyset \neq L^+(x) \subset A \}
\]

is the region of attraction (or stable manifold) of \( A \). If \( W^+(A) \) is an open neighbourhood of \( A \), then \( A \) is called an attractor.

A compact invariant set \( A \subset M \) is called stable (in the sense of Lyapunov) if every neighbourhood of \( A \) contains a positively invariant neighbourhood of \( A \). This is equivalent to saying that \( \emptyset \neq J^+(x) \subset A \) for every \( x \in A \) (see [3, Theorem 1.12]). A stable attractor \( A \) is also called asymptotically stable. In this case \( \emptyset \neq J^+(x) \subset A \) for every \( x \in W^+(A) \) and the restricted flow in \( W^+(A) \setminus A \) is particularly simple, since it is parallelizable with a compact global section (see [3, p. 83]). If \( A \) is an attractor in \( M \), it is well-known that the positive first prolongation

\[
D^+(A) = \bigcup_{x \in A} D^+(x)
\]

of \( A \) is compact invariant and asymptotically stable with \( W^+(D^+(A)) = W^+(A) \) (see [2, Theorem 8.20]). Note that \( D^+(A) = \{ x \in M : L^-(x) \cap A \neq \emptyset \} \), if \( A \) is an attractor.

Let \( A \subset M \) be an isolated compact invariant set and let \((N, N^+, N^-)\) be an isolating block of \( A \). The final entrance time function \( f : W^+(A) \setminus A \rightarrow \mathbb{R} \)
Lemma 3.1. The final entrance time function $f$ is discontinuous at $x \in W^+(A) \setminus A$ if and only if there are $x_n \to x$ such that $f(x_n) \to +\infty$.

Proof. If $f$ is discontinuous at $x$, there are $x_n \to x$, $x_n \in W^+(A) \setminus A$, and $\epsilon > 0$ such that $f(x_n) > f(x) + \epsilon$, for every $n \in \mathbb{N}$. By the continuity of the flow and passing to a subsequence if necessary, there are $s_n \to f(x)$ such that $x_n \in N^+ \setminus N^-$, since $f(x)x \in \alpha^+$. Thus, eventually $f(x_n) > f(x) + \epsilon > s_n$. This means that $C^+(s_n, x_n)$ exits $N$ eventually for all large enough $n \in \mathbb{N}$. The exit time function $\sigma^+: N \to [0, +\infty]$ defined by $\sigma^+(x) = 0$, if $x \in N^-$, and

$$\sigma^+(x) = \sup\{t > 0 : [0, t)x \in N \setminus N^-,\}$$

if $x \in N \setminus N^-$, is continuous. Hence, $f(x_n) > s_n + \sigma^+(s_n, x_n) \to +\infty$, since $\sigma^+(f(x)x) = +\infty$. The converse is obvious.

Proposition 3.2. Let $A$ be an isolated invariant set of a flow $\phi$ on a separable, locally compact, metrizable space $M$ and let $(N, N^+, N^-)$ be an isolating block of $A$. Then:

(i) $\phi$ maps $\mathbb{R} \times \text{int}_{\partial N} \alpha^+$ homeomorphically onto $\text{int}_{\partial N} \alpha^+$, and

(ii) $\phi$ maps $\mathbb{R} \times \partial_{\partial N} \alpha^+$ homeomorphically onto $\partial_{\partial N} \alpha^+$.

Proof. It is clear that $\phi$ maps $\mathbb{R} \times \alpha^+$ in a one-to-one manner onto $\mathbb{R} \alpha^+$. If now $f : W^+(A) \setminus A \to \mathbb{R}$ is the final entrance time function of the isolating block, then $f|\text{int}_{\partial N} \alpha^+$ is continuous, by Lemmas 2.1 and 3.1. The inverse of $\phi|\mathbb{R} \alpha^+$ is the map $h : \mathbb{R} \alpha^+ \to \mathbb{R} \times \alpha^+$ defined by $h(x) = (-f(x), f(x)x)$, which is therefore continuous on $\text{int}_{\partial N} \alpha^+$. To prove the second assertion, suppose that $\phi$ does not map $\mathbb{R} \times \partial_{\partial N} \alpha^+$ homeomorphically onto $\partial_{\partial N} \alpha^+$. This means that there are $t, t_n \in \mathbb{R}$ and $x, x_n \in \partial_{\partial N} \alpha^+$, $n \in \mathbb{N}$, such that $t_n x_n \to tx$, but the sequence $\{t_n, x_n : n \in \mathbb{N}\}$ does not converge to $(t, x)$. Suppose first that $\{t_n : n \in \mathbb{N}\}$ is bounded. If $s \in \mathbb{R}$ is any limit point of $\{t_n : n \in \mathbb{N}\}$ and $y \in \partial_{\partial N} \alpha^+$ is any limit point of $\{x_n : n \in \mathbb{N}\}$, then $tx = sy$ and therefore $t = s$ and $x = y$. Since $\alpha^+$ is compact, it follows that $t_n \to t$ and $x_n \to x$, contradiction. So $\{t_n : n \in \mathbb{N}\}$ is unbounded and passing to a subsequence if necessary we may assume that $t_n \to +\infty$. For every $s \in \mathbb{R}$ we have eventually $t_n - t + s > 0$ for large values of $n \in \mathbb{N}$. So, $(t_n - t + s)x_n \in N$ and $(t_n - t + s)x_n \to sx$. Hence $sx \in N$ for every $s \in \mathbb{R}$, which contradicts the fact that $A$ is isolated.

Corollary 3.3. Let $\phi$ be a continuous flow on a separable, locally compact, metrizable space $M$ and let $A \subset M$ be an invariant continuum which is an
isolated unstable attractor with region of attraction $W^+(A)$. Let $(N,N^+,N^-)$ be an isolating block of $A$ such that $N \subset W^+(A)$. Then,

$$\mathbb{R} \partial_D N \alpha^+ = \{ x \in W^+(A) \setminus A : J^+(x) \notin A \}$$

and is homeomorphic image of $\mathbb{R} \times \partial_D N \alpha^+$ under $\phi$.

Concerning the existence of explosions in the region of attraction of an isolated unstable attractor, we have the following:

**Theorem 3.4.** Let $A$ be an invariant continuum which is an isolated unstable attractor of a continuous flow on a connected, locally compact, metrizable space $M$. If $W^+(A) \setminus A$ is connected, then either $D^+(A) = M$ or there exists a point $x \in W^+(A) \setminus D^+(A)$ such that $\emptyset \neq J^+(x) \not\subset A$.

**Proof.** Suppose that the conclusion is not true. Since $D^+(A) \neq M$, we have $D^+(A) \neq W^+(A)$, because $A$ is an attractor and $M$ is assumed to be connected. The restricted flow in $W^+(A) \setminus D^+(A)$ is parallelizable and has a compact global section $S$. Let $(N,N^+,N^-)$ be an isolating block of $A$ such that $N \subset D^+(A) \cup (0,\infty)S$. By Lemma 3.1 and since we suppose that $\emptyset \neq J^+(x) \subset A$ for every $x \in W^+(A) \setminus D^+(A)$, the final entrance time function $f$ is continuous on $W^+(A) \setminus D^+(A)$ and hence bounded on $S$. If $W^+(A) \setminus D^+(A)$ is closed in $W^+(A) \setminus A$, then we arrive at a contradiction, since it is already open and $W^+(A) \setminus A$ is connected by assumption, which implies that $D^+(A) = A$, that is $A$ would be stable. So $W^+(A) \setminus D^+(A)$ is not closed in $W^+(A) \setminus A$, which means that there exists a sequence $\{x_n : n \in \mathbb{N}\}$ of points of $W^+(A) \setminus D^+(A)$ converging to some point $x \in D^+(A) \setminus A$. There are eventually $t_n < 0$ such that $t_n x_n \in S$, for every $n \in \mathbb{N}$, and since $S$ is compact, we may assume that $t_n x_n \to y$ for some $y \in S$, passing to a subsequence. Obviously, $f(t_n x_n) = f(x_n) - t_n$. If $\{t_n : n \in \mathbb{N}\}$ had a bounded subsequence, and hence some convergent, we would have $y \in C(x)$, contradiction. So, we must have $t_n \to -\infty$. But then $x \in J^+(y)$, which contradicts our assumption in the beginning that the conclusion is not true.

The case $D^+(A) = M$ may occur as the example of the flow on $S^1$ with a single singular point shows. Moreover, in this example there is no explosion in $M \setminus A$. As an illustration of Theorem 3.4 consider the smooth flow on $\mathbb{R}^2$ defined by the differential equation (in polar coordinates)

$$r' = r(1 - r) \quad \text{and} \quad \theta' = \sin^2 \left( \frac{\theta}{2} \right).$$

Then, $\{(1,0)\}$ is an isolated unstable attractor with $W^+(1,0) = \mathbb{R}^2 \setminus \{(0,0)\}$ and $D^+(1,0) = S^1$. For $0 < s < 1$ or $s > 1$ the orbit of the point $(s,0) \in W^+(1,0) \setminus D^+(1,0)$ explodes, because $J^+(s,0) = S^1 = D^+(1,0)$. By the way, the closed disc of radius $1/2$ centered at $(1,0)$ is an isolating block and the flow does not map $\mathbb{R} \times \alpha^+$ homeomorphically onto $\mathbb{R} \alpha^+$. 
If $A$ is an isolated unstable attractor the question arises whether the flow explodes at orbits in $D^+(A) \setminus A$. In the example on $\mathbb{R}^2$ of differential Equation (1) we have $A = \{(1,0)\}$, $D^+(A) = J^+(1,0) = S^1$ and the disc $N$ of radius 1/2 centered at $(1,0)$ is an isolating block. In this case the flow does not explode at the points of $D^+(A) \setminus A$, because $D^+(A) \cap \alpha^+ \subseteq \text{int}_N \alpha^+$. On the other hand, there is a flow on the 2-torus with a single fixed point $x_0$, which is an isolated unstable attractor with $D^+(x_0) = T^2$ and the flow explodes at two orbits. This flow is induced on $T^2$ from the differential equation (in polar coordinates)

$$r' = (r - 1)(2 - r) \quad \text{and} \quad \theta' = \sin^2 \left( \frac{\theta}{2} \right)$$

on the annulus $1 \leq r \leq 2$ and $0 \leq \theta \leq 2\pi$.

A sufficient condition for explosion in $D^+(A) \setminus A$ is given by the following:

**Theorem 3.5.** Let $A$ be an invariant continuum which is an isolated unstable attractor of a flow on a locally compact, metrizable space $M$. If the first integral Alexander-Spanier cohomology group $\overline{H}^1(W^+(A); \mathbb{Z})$ is trivial, then there is a point $x \in D^+(A) \setminus A$ such that $J^+(x) \not\subseteq A$ and therefore the flow explodes at some orbit in $D^+(A) \setminus A$.

**Proof.** Suppose that $\emptyset \neq J^+(x) \subseteq A$ for every $x \in D^+(A) \setminus A$. Let $\chi : \mathbb{R} \to (0,1)$ be an increasing homeomorphism. Let $(N,N^+,N^-)$ be an isolating block of $A$ such that $N \subset W^+(A)$ and let $f : W^+(A) \setminus A \to \mathbb{R}$ be the final entrance time function. Let $g : D^+(A) \to S^1$ be the function defined by

$$g(x) = \begin{cases} 1, & \text{if } x \in A, \\ \exp(2\pi i \chi(f(x))), & \text{if } x \in D^+(A) \setminus A. \end{cases}$$

Our assumption combined with Lemma 3.1 implies that $g$ is continuous. Since $D^+(A)$ is an asymptotically stable invariant continuum with $W^+(D^+(A)) = W^+(A)$, the restricted flow in $W^+(A) \setminus D^+(A)$ is parallelizable with a compact global section $S$. The set $W_0 = D^+(A) \cup [0, +\infty) S$ is a positively invariant, compact, strong deformation retract of $W^+(A)$. If $W_n = D^+(A) \cup [n, +\infty) S$, $n \in \mathbb{N}$, then $W_n$ is a strong deformation retract of $W_m$ for $n \geq m$, and so the inclusion $W_n \subset W_m$ induces an isomorphism in Alexander-Spanier cohomology. Since

$$D^+(A) = \bigcap_{n=0}^{\infty} W_n,$$

it follows from the continuity property of the Alexander-Spanier cohomology (see [12, p. 318]) that $\overline{H}^1(W^+(A); \mathbb{Z}) \cong \overline{H}^1(D^+(A); \mathbb{Z})$, and so the latter is trivial. Recall that $\overline{H}^1(D^+(A); \mathbb{Z})$ is naturally isomorphic to the abelian group of homotopy classes of continuous maps of $D^+(A)$ to $S^1$ [7, Theorem
Thus, \( g \) must be homotopic to a constant, that is \( g = \exp(2\pi ih) \), for some continuous function \( h : D^+(A) \to \mathbb{R} \). Since \( A \) is connected, \( h \) takes an integer value, say \( k \), on \( A \). Moreover, for every \( x \in D^+(A) \setminus A \), the function \( \psi : C^+(x) \cup A \to \mathbb{Z} \) defined by \( \psi(tx) = h(tx) - \chi(f(tx)) \), for \( t \geq 0 \), and \( \psi(A) = k \), is continuous and hence constant, because \( \emptyset \neq L^+(x) \subset A \). This implies that \( h(x) = \chi(f(x)) + k \), for every \( x \in D^+(A) \setminus A \). On the other hand, \( L^-(x) \cap A \neq \emptyset \), for every \( x \in D^+(A) \setminus A \), which means that there are \( t_n \to -\infty \), such that the sequence \( \{t_n \in \mathbb{N} \} \) converges to some point of \( A \). Consequently, \( f(t_n x) \to +\infty \) and \( k = h(A) = \lim_{n \to +\infty} h(t_n x) = 1 + k \), contradiction.

If \( A \) is not isolated, Theorem 3.5 is not true. Consider for example the extension to the 2-sphere \( S^2 \) of the parallel flow on \( \mathbb{R}^2 \). The point at infinity \( \infty \) is fixed and is an unstable attractor, but not isolated. Its region of attraction is \( W^+(\infty) = S^2 = D^+(\infty) \) and \( J^+(x) = \{\infty\} \) for every \( x \in S^2 \setminus \{\infty\} \). A more interesting example is the following: In [1, Section 4] a continuous flow on the 4-sphere \( S^4 \) is constructed, in which the dyadic solenoid \( Y = \{(z_n)_{n \geq 0} : z_n \in S^1 \text{ and } z_n = z_{n+1}^2, n \geq 0 \} \) embeds as an isolated minimal invariant set. Considering \( S^4 \) as the unreduced suspension of the 3-sphere \( S^3 \), that is the quotient space obtained from \( S^3 \times [-1, 1] \) by identifying \( S^3 \times \{\pm 1\} \) to points, the poles of \( S^4 \), \( Y \) embeds to \( \{[y, 0] : y \in Y \} \) and the flow has the following properties:

(i) The poles of \( S^4 \) are fixed points.
(ii) If \( y \notin Y \) and \( s \neq \pm 1 \), then \( J^+[y, s] \) is the north pole and \( J^-[y, s] \) is the south pole.
(iii) If \( y \in Y \) and \( 0 < s < 1 \), then \( J^+[y, s] \) is the north pole and \( L^-[y, s] = Y \), while if \( -1 < s < 0 \), then \( J^-[y, s] \) is the south pole and \( L^+[y, s] = Y \).

If \( y_0 \in Y \), the invariant continuum \( A = Y \cup \overline{C[y_0, 1/2]} \) is a non-isolated unstable attractor with \( D^+(A) = Y \cup \bigcup_{y \in Y} \overline{C[y, 1/2]} \) and \( W^+(A) = S^4 \setminus \{\text{south pole}\} \). So \( \overline{H}^1(W^+(A); \mathbb{Z}) \) is trivial but \( J^+[y, s] = \{\text{north pole}\} \subset A \) for every \( [y, s] \in D^+(A) \setminus A \). In this example \( \overline{H}^1(A; \mathbb{Z}) \cong \overline{H}^1(Y; \mathbb{Z}) \) is also not finitely generated.

We shall examine now topological properties of an isolated unstable attractor \( A \) in case there is essentially no explosion in \( D^+(A) \setminus A \) and \( D^+(A) \setminus A \) is not too bad.

**Proposition 3.6.** Let \( A \) be an invariant continuum which is an isolated unstable attractor of a flow on a locally compact, metrizable space \( M \). If \( D^+(A) \setminus A \) is locally connected, then it has a finite number of connected components. If moreover \( \emptyset \neq J^+(x) \subset A \) for every \( x \in D^+(A) \setminus A \), then the number of the connected components of \( D^+(A) \setminus A \) is not greater than \( \text{rank} \overline{H}^1(W^+(A); \mathbb{Z}) \).
Proof. We shall prove first that \( D^+(A) \setminus A \) has a finite number of connected components. Every connected component of \( D^+(A) \setminus A \) is open in \( D^+(A) \), since \( D^+(A) \setminus A \) is assumed to be locally connected. Suppose that \( D^+(A) \setminus A \) has an infinite number of connected components \( C_n, n \in \mathbb{N} \). Let \( (N,N^+,N^-) \) be an isolating block of \( A \). Then, \( C_n \cap \partial N \neq \emptyset \) and if we pick points \( x_n \in C_n \cap \partial N, n \in \mathbb{N} \), they have some limit point \( x \in D^+(A) \cap \partial N \subset D^+(A) \setminus A \), which must therefore be in some connected component \( C \) of \( D^+(A) \setminus A \). But then \( C \) is an open neighborhood of \( x \) in \( D^+(A) \) and so we must have \( x_n \in C \) for infinitely many values of \( n \), contradiction. Thus, \( D^+(A) \setminus A \) has a finite number of connected components, say \( C_1,\ldots,C_k \). To prove the second assertion, note first that \( C_j \cup A \) is an invariant continuum for all \( j = 1,2,\ldots,k \). Let \( f : W^+(A) \setminus A \to \mathbb{R} \) be the final entrance time function of the isolating block. For each \( j = 1,2,\ldots,k \) let \( g_j : D^+(A) \to S^1 \) be the function defined by

\[
g_j(x) = \begin{cases} 
1, & \text{if } x \in D^+(A) \setminus C_j, \\
\exp(2\pi i \chi(f(x))), & \text{if } x \in C_j,
\end{cases}
\]

where \( \chi : \mathbb{R} \to (0,1) \) is an increasing homeomorphism. Since \( D^+(A) \setminus A \) has finitely many connected components and \( f \) is continuous on \( D^+(A) \setminus A \), \( g_j \) is continuous. The argument in the proof of Theorem 3.5 shows that \( g_j|C_j \cup A \) is not homotopic to a constant. Consequently, \( (g_j|C_j \cup A)^m \), for \( m \neq 0 \), defines a nonzero element of \( \overline{H}^1(C_j \cup A; \mathbb{Z}) \). We shall prove that \( g_1,\ldots,g_k \) define a set of linearly independent cohomology classes in \( \overline{H}^1(D^+(A); \mathbb{Z}) \). Indeed, let \( n_1,\ldots,n_k \in \mathbb{Z} \) be such that \( g_1^{n_1} \ldots g_k^{n_k} \) is homotopic to a constant. There is a continuous function \( h : D^+(A) \to \mathbb{R} \) such that \( g_1^{n_1} \ldots g_k^{n_k} = \exp(2\pi ih) \). It follows from the definitions that \( (g_j|C_j \cup A)^{n_j} = \exp(2\pi i(h|C_j \cup A)) \), and so it is homotopic to a constant. Hence \( n_j = 0 \), for every \( j = 1,2,\ldots,k \). This completes the proof.

Theorem 3.7. Let \( A \) be an invariant continuum and an isolated unstable attractor of a flow \( \phi \) on a locally compact ANR \( M \), such that:

(i) \( D^+(A) \setminus A \) is an ANR, and
(ii) \( \emptyset \neq J^+(x) \subset A \) for every \( x \in D^+(A) \setminus A \).

Then \( \overline{H}^q(A; \mathbb{Z}) \) is finitely generated for all \( q \).

Proof. We observe first that the restricted flow on the invariant locally compact subspace \( D^+(A) \setminus A \) of \( M \) is parallelizable with a compact global section. Actually, if \( (N,N^+,N^-) \) is an isolating block of \( A \), such that \( N \subset W^+(A) \), then \( S = D^+(A) \cap \alpha^+ \subset \text{int}_N \alpha^+ \), by Assumption (ii) and Lemma 2.1, and is a compact global section to the restricted flow on \( D^+(A) \setminus A \), by Proposition 3.2. The assumption that \( D^+(A) \setminus A \) is an ANR implies that \( S \) is an ANR also. It follows that \( H^q(S; \mathbb{Z}) \) is finitely generated for all \( q \) and we
have the chain of isomorphisms
\[ H^{q+1}_c(D^+ - \{x\}, \mathbb{Z}) \cong H^{q+1}_c(D^+ \setminus \{x\}; \mathbb{R} \times S; \mathbb{Z}) \cong H^q(S; \mathbb{Z}) \cong H^q(S; \mathbb{Z}), \]
where \( H^*_c \) denotes Alexander-Spanier cohomology with compact supports (see [12, p. 320]). From the exact sequence of the pair \((D^+ - \{x\}, \mathbb{R} \times S; \mathbb{Z})\) we deduce that \( H^q(A; \mathbb{Z}) \) is finitely generated for all \( q \).

4. Singular unstable attractors in the 2-sphere.

Let \( \phi \) be a continuous flow on a connected 2-manifold \( M \) having a fixed point \( x_0 \), which we assume to be an isolated invariant set. Every neighbourhood of \( x_0 \) contains an isolating compact neighbourhood of \( x_0 \), which is a compact 2-manifold with boundary. The interior \( V \) of such a neighbourhood is an open 2-manifold of finite genus and with finitely many ends. We can reparametrize the local flow in \( V \) to get a flow with respect to which \( \{x_0\} \) is the only invariant set in \( V \), apart \( V \) itself. It follows from the Smoothing Theorem of C. Gutierrez (see [10]) that the flow in \( V \) is topologically equivalent to a smooth flow, that is there exists a homeomorphism \( h: V \to V \) which maps the oriented orbits in \( V \) onto the oriented orbits of some smooth flow in \( V \). Since \( \{x_0\} \) is an isolated invariant set with respect to the smooth flow also, it has a smooth isolated block \((P, P^+, P^-)\) with respect to the smooth flow. Clearly, \((h^{-1}(P), h^{-1}(P^+), h^{-1}(P^-))\) is an isolating block for \( \{x_0\} \) with respect to the initial flow \( \phi \), and \( h^{-1}(P) \) is a topological compact 2-manifold with boundary \( h^{-1}(P^+) \cup h^{-1}(P^-) \). The sets \( h^{-1}(P^+), h^{-1}(P^-) \) are topological compact 1-manifolds with common boundary \( h^{-1}(P^+) \cap h^{-1}(P^-) \), which is a finite set. If we start with a disc neighbourhood, then \( V \) has genus zero and so does \( h^{-1}(P) \). Summarizing we have the following:

**Lemma 4.1.** Let \( \phi \) be a continuous flow on a connected 2-manifold \( M \) and let \( x_0 \in M \) be a fixed point of the flow which is an isolated invariant set. Then every neighbourhood of \( x_0 \) contains an isolating block \((N, N^+, N^-)\) such that \( N \) is a genus zero compact 2-manifold with boundary, and \( N^+, N^- \) are compact 1-manifolds with common boundary \( N^+ \cap N^- \), which is a finite set.

**Proposition 4.2.** The set \( \partial_{\partial N} \alpha^+ \) is finite and \( \alpha^+ \) has finitely many connected components.

**Proof.** Suppose on the contrary that \( \partial_{\partial N} \alpha^+ \) has infinitely many points. By compactness, there is a sequence \( \{x_n : n \in \mathbb{N}\} \) in \( \partial_{\partial N} \alpha^+ \) converging to a point \( x \in \partial_{\partial N} \alpha^+ \). Let \( I \) be a small open interval in \( N^+ \setminus N^+ \cap N^- \) containing \( x \). Passing to a subsequence if necessary, we may assume that \( x_n \in I \) and if \([x_n, x]\) is the subinterval of \( I \) with endpoints \( x_n \) and \( x \), then
Let \([x_{n+1}, x] \subset [x_n, x]\), for every \(n \in \mathbb{N}\). Let \(U\) be a disc neighbourhood of \(x_0\) contained in \(\text{int} N\). There is a \(T > 0\) such that \(C^+(Tx) \subset U\). A similar argument as in the proof of Lemma 2.1(a) shows that we must also have \(C^+(Tx_n) \subset U\), for large \(n\). Since the flow is transverse to \(N^+ \setminus N^+ \cap N^-\) into \(N\), some small closed interval \(J \subset I\) containing \(x\) in its interior is a local section to the flow of extent \(T\), by [11, Lemma VII.2.9]. This means that the flow maps \([-T, T] \times J\) homeomorphically onto its image. So, \(\phi([0, T] \times J)\) is a disc in \(N\) and \(J = \phi([0, T] \times J) \cap \partial N\). For large \(n\) we have \(x_n \in J\). The simple closed curve \(T[x_n, x] \cup C^+(Tx_n) \cup C^+(Tx) \cup \{x_0\}\) bounds a disc \(V_n\) in \(U\), by the Jordan-Schönflies theorem. It follows that the set \(D_n = \phi([0, T] \times J) \cup V_n\) is a disc in \(N\) and \(\partial D_n = [x_n, x] \cup C^+(x_n) \cup C^+(x) \cup \{x_0\}\). It is evident that \(D_n\) is positively invariant, since the flow is transverse to \([x_n, x]\) into \(D_n\) and no positive semiorbit starting on \((x_n, x)\) can cross \(\partial D_n\). This implies that \([x_n, x] \subset \alpha^+\) and hence \(x_{n+1} \in \text{int} \partial N \alpha^+\). This contradiction proves the first assertion. The second follows from the first.

In Proposition 4.2 we have not assumed that the fixed point is an attractor. Combining it now with Corollary 3.3 we get the following:

**Corollary 4.3.** Let \(x_0\) be a fixed point of a continuous flow on a connected 2-manifold \(M\). If \(x_0\) is an isolated unstable attractor, then the flow on \(W^+(x_0) \setminus \{x_0\}\) explodes at finitely many orbits.

In the rest of this section we shall be concerned with the structure of the flow in the region of attraction of a fixed point of a flow on the 2-sphere \(S^2\), which is an unstable attractor. We are mainly interested in the topological structure of the positive first prolongation of the fixed point and the dynamics of the flow in it. Most of the results will be proved without the assumption that the fixed point is an isolated invariant set.

**Proposition 4.4.** Let \(x_0 \in S^2\) be a fixed point of a continuous flow \(\phi\) on \(S^2\). If \(x_0\) is an unstable attractor, then \(L^+(x) = L^-(x) = \{x_0\}\) for every \(x \in D^+(x_0)\).

**Proof.** Since \(D^+(x_0)\) is a compact invariant set, we have \(\emptyset \neq L^-(x) \subset D^+(x_0)\) for every \(x \in D^+(x_0)\). If \(y \in L^-(x)\), then \(x_0 \in L^+(y) \subset L^-(x)\). Suppose that \(L^-(x)\) contains and some other point \(z \neq x_0\). Then, \(z \in D^+(x_0)\) and \(z\) is not fixed. Moreover, \(L^+(z) = L^-(z) = \{x_0\}\), because \(x_0\) is the only fixed point in \(W^+(x_0)\) (see [11, Proposition 1.11]). Thus, \(\overline{C(z)}\) is a simple closed curve and \(S^2 \setminus \overline{C(z)}\) has two connected components with common boundary \(C(z)\), whose closures are closed discs by the Jordan-Schönflies theorem. Both are invariant and \(C(x)\) is contained in one of them, call it \(B\). Let \(S\) be a local section to the flow at \(z\). That is, \(S\) is an open arc passing through \(z\) and \(\phi\) maps \((-\epsilon, \epsilon) \times S\) homeomorphically onto an open neighbourhood of \(z\) for some \(\epsilon > 0\) (see [11, Corollary 2.6]). Since \(z \in L^-(x)\), there is some \(t < 0\) such that \(tx \in S\) and we can choose \(t\)
such that $C^+(tx) \cap S = \{tx\}$, because $L^+(x) = \{x_0\}$. Let $J$ be the open subarc in $S$ with endpoints $tx$ and $z$. The set $K = \overline{C^+(tx)} \cup J \cup \overline{C^+(z)}$ is a simple closed curve and $S^2 \setminus K$ has two connected components. One of them is positively invariant and the other is negatively invariant. Let $D$ be the one that is negatively invariant. Then $(-\infty,0)x \subset D$. Since $z \in L^-(x)$ and $J \subset B$, there is some $s < t$ such that $sx \in J$. But then $tx \in (0, +\infty)(sx) \subset S^2 \setminus \overline{D} = S^2 \setminus J$, contradiction.

**Theorem 4.5.** Let $x_0$ be a fixed point of a continuous flow $\phi$ on $S^2$. If $x_0$ is an unstable attractor, then every point $x \in D^+(x_0) \setminus \{x_0\}$ has an open neighbourhood $V$ such that $(D^+(x_0) \setminus \{x_0\}) \cap V$ is homeomorphic to the open interval $(-1,1)$ or the open rectangle $(-1,1) \times (-1,1)$.

**Proof.** Let $x \in D^+(x_0) \setminus \{x_0\}$. Since $x$ is not fixed, there is a local section to the flow through $x$. Thus there are open arcs $S$ containing $x$ and $\epsilon > 0$ such that $\phi$ maps $(-\epsilon,\epsilon) \times S$ homeomorphically onto an open neighbourhood of $x$. We can shrink $S$, if necessary, so that $C(x) \cap S = \{x\}$, because $L^+(x) = L^-(x) = \{x_0\}$, by Proposition 4.4. The set $S \setminus \{x\}$ is the disjoint union of two open intervals $S_1, S_2$ in $S$ with one common endpoint $x$. If there are open intervals $I \subset S_1$ and $J \subset S_2$ with common endpoint $x$ such that $I \cup J \subset W^+(x_0) \setminus D^+(x_0)$, or $I \cup J \subset D^+(x_0)$, or $I \subset W^+(x_0) \setminus D^+(x_0)$ and $J \subset D^+(x_0)$, or vice versa, then we are done. We shall prove that otherwise we are led to a contradiction. Suppose that every open interval $I \subset S_1$ with one endpoint $x$ contains points of $D^+(x_0)$ and $W^+(x_0) \setminus D^+(x_0)$. There is a sequence $\{x_n : n \in \mathbb{N}\}$ of points in $D^+(x_0) \cap S_1$, which converges monotonically to $x$, such that $[x_n, x_{n+1}] \cap (W^+(x_0) \setminus D^+(x_0)) \neq \emptyset$, where $[x_n, x_{n+1}]$ denotes the closed interval on $S_1$ with endpoints $x_n$ and $x_{n+1}$. The monotonicity of convergence to $x$ means that $[x, x_{n+1}] \subset [x, x_n]$ for every $n \in \mathbb{N}$. From Proposition 4.4 we have $L^+(x_n) = L^-(x_n) = \{x_0\}$ for every $n \in \mathbb{N}$. Let $D$ be the closed invariant disc in $S^2$ bounded by the simple closed curve $\overline{C(x)}$, such that $[x, x_n] \subset D$ for some (and hence for all) $n \in \mathbb{N}$. Then, $\overline{C(x_n)}$ bounds a closed disc $D_n \subset D$ such that $[x, x_n] \subset D \setminus \text{int}D_n$. The monotonicity implies that $D_n \subset D_m$ and $[x_n, x_m] \subset D_m \setminus \text{int}D_n$ for $n < m$. If $y_n \in [x_n, x_{n+1}] \cap (W^+(x_0) \setminus D^+(x_0))$, then $C(y_n) \subset \text{int}D_{n+1} \setminus D_n$ and it hits some connected component of a global section $\Sigma$ of the parallelizable restricted flow in $W^+(x_0) \setminus D^+(x_0)$, which is a compact 1-manifold, that is $\Sigma$ is finite union of disjoint simple closed curves $\Sigma_1, \ldots, \Sigma_k$. If $C(y_n) \cap \Sigma_j \neq \emptyset$, then $\Sigma_j \subset \text{int}D_{n+1} \setminus D_n$, because $\Sigma \cap \overline{C(x_{n+1})} \cup \overline{C(x_n)} = \emptyset$. But the sets $\text{int}D_{n+1} \setminus D_n, n \in \mathbb{N}$ are disjoint and infinitely many. This contradiction completes the proof.

From Proposition 3.6 and Theorem 4.5 we have the following:

**Corollary 4.6.** If $x_0 \in S^2$ is a fixed point of a continuous flow on $S^2$ and is an isolated unstable attractor, then $D^+(x_0) \setminus \{x_0\}$ is an ANR and has
Theorem 4.7. If \( t \) and is non-wandering with respect to the restricted flow in \( D \) an unstable attractor, then the restricted flow in \( D \) continuous flow on \( S \) and noncompact boundary, with finitely many boundary components.

The dynamics in the positive first prolongation of a fixed point of a continuous flow on \( S^2 \) which is an unstable attractor is described by the following:

**Theorem 4.7.** If \( x_0 \in S^2 \) is a fixed point of a continuous flow on \( S^2 \) and is an unstable attractor, then the restricted flow in \( D^+(x_0) \setminus \{x_0\} \) is completely unstable.

**Proof.** Suppose on the contrary that there is some \( x \in D^+(x_0) \setminus \{x_0\} \), which is non-wandering with respect to the restricted flow in \( D^+(x_0) \setminus \{x_0\} \). This means that there are a sequence \( \{x_n : n \in \mathbb{N}\} \) of points of \( D^+(x_0) \setminus \{x_0\} \) and \( t_n \rightarrow +\infty \) such that \( x_n \rightarrow x \) and \( t_n x_n \rightarrow x \). By Proposition 4.4, \( C(x) \) is a simple closed curve and bounds an invariant closed disc \( D \) in \( S^2 \) such that \( C(x_n) \subset intD \) for every \( n \in \mathbb{N} \), passing to a subsequence if necessary. Let \( S \) be an open arc that is a local section to the flow at \( x \) of some extent \( \epsilon > 0 \), and such that \( C(x) \cap S = \{x\} \). Let \( I \) be the connected component of \( S \setminus \{x\} \) which lies in \( intD \). For large enough values of \( n \), there are \( |s_n| < \epsilon \) and \( |r_n| < \epsilon \) such that \( s_n x_n \in I \) and \( (t_n + r_n)x_n \in I \). Since \( t_n \rightarrow +\infty \), eventually we have \( t_n + r_n > s_n + \epsilon \). Let \( T_n > \epsilon \) be the first time \( C(s_n x_n) \) hits \( I \). If \( [s_n x_n, T_n x_n] \) denotes the closed interval in \( I \) with endpoints \( s_n x_n \) and \( T_n x_n \), then \( K = [s_n, T_n] x_n \cup [s_n x_n, T_n x_n] \) is a simple closed curve and bounds a closed disc \( D' \subset intD \), which is positively or negatively invariant. But then \( L^-(x_n) \subset D' \) or \( L^+(x_n) \subset D' \), which contradicts that \( x_n \in D^+(x_0) \setminus \{x_0\} \).

Recall that a separatrix in a completely unstable flow is an orbit with nonempty positive prolongational limit set. It follows from Corollary 4.3 that if \( x_0 \) is a fixed point of a continuous flow on \( S^2 \) which is an isolated unstable attractor, then the restricted flow in \( D^+(x_0) \setminus \{x_0\} \) is completely unstable with finitely many separatrices.

**References**


Received July 21, 2000 and revised May 21, 2001.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CRETE
GR-71409 IRAKLION
GREECE
E-mail address: athanako@math.uoc.gr
ASYMPTOTIC BEHAVIOR OF POSITIVE SOLUTIONS OF SOME ELLIPTIC PROBLEMS

E.N. Dancer, Yihong Du, and Li Ma

In this paper, we discuss the asymptotic behavior of the positive solutions of the problem
\[-\Delta u = au - bu^p, \ u|_{\partial \Omega} = 0\]
as \(p \to 1 + 0\) and as \(p \to \infty\). We show that, for each case, the behavior is determined by a limiting problem. Moreover, the limiting problem is of free boundary nature when \(p \to \infty\).

1. Introduction and main results.

In this paper, we study the asymptotic behavior of positive solutions of the problem
\[(1.1) \quad -\Delta u = au - b(x)u^p, \ x \in \Omega; \ u = 0, \ x \in \partial \Omega,\]
for \(p\) near 1 and near \(\infty\), respectively. Here \(\Omega\) is a bounded smooth domain in \(\mathbb{R}^N\) \((N \geq 1)\) and \(b(x)\) is a nonnegative function in \(C(\overline{\Omega})\), \(a\) and \(p\) are constants but the exponent \(p\) is always greater than 1.

Problem (1.1) arises from mathematical biology and Riemannian geometry, and has attracted considerable interests; see, for example, \([AT], [AM], [D], [dP], [DH], [FKLM], [KW], [M]\) and \([O]\). The dependence of the positive solutions of (1.1) on the parameter \(a\) is well understood but little is known about the dependence on \(p\).

When \(b(x)\) is strictly positive on \(\Omega\), (1.1) is the steady-state logistic equation and it is well-known that for fixed \(p > 1\) it has no positive solution if \(a \leq \lambda_1^{\Omega}\) and there is a unique positive solution \(u = u_a\) when \(a > \lambda_1^{\Omega}\), where \(\lambda_1^{\Omega}\) denotes the first eigenvalue of the problem
\[-\Delta u = \lambda u, \ u|_{\partial \Omega} = 0.\]

Moreover, \(a \to u_a\) is continuous and strictly increasing as a function from \((\lambda_1^{\Omega}, \infty)\) to \(C(\overline{\Omega})\) (with the natural order), and
\[
\lim_{a \to \lambda_1^{\Omega} + 0} u_a(x) = 0 \text{ uniformly in } \overline{\Omega};
\]
\[
\lim_{a \to \infty} u_a(x) = \infty \text{ uniformly on any compact subset of } \Omega.
\]

When \(b^{-1}(0) := \{x \in \Omega : b(x) = 0\}\) is a proper subset of \(\Omega\), the behavior of (1.1) is more complicated. Assume for simplicity that \(b^{-1}(0) = \overline{\Omega}_0 \subset \subset \Omega\),

215
where $\Omega_0$ is open, connected and with smooth boundary. Then it is well-known that (1.1) has no positive solution unless $a \in (\lambda_1^\Omega, \lambda_1^{\Omega_0})$, in which case there is a unique positive solution $u_a$ which varies continuously with $a$ and is strictly increasing in $a$. Moreover, $u_a \to 0$ uniformly on $\Omega$ as $a \to \lambda_1^\Omega + 0$, but as $a \to \lambda_1^{\Omega_0}$, $u_a(x) \to \infty$ uniformly on $\Omega_0$ and $u_a \to U$ uniformly on any compact subset of $\Omega \setminus \Omega_0$, where $U$ is the unique minimal positive solution of the boundary blow-up problem

$$-\Delta u = au - b(x)u^p, \quad x \in \Omega \setminus \Omega_0; \quad u|_{\partial \Omega} = 0, \quad u|_{\partial \Omega_0} = \infty.$$  

We refer to [DH] and the references therein for more details.

To understand the effect of the exponent $p$ on the unique positive solution of (1.1), we fix $a$ and consider the extreme cases, that is when $p \to 1 + 0$ and when $p \to \infty$. In each case, we obtain a limiting problem which determines the asymptotical behavior of (1.1).

To describe our results, we need to recall several simple properties of the first eigenvalue of the Laplacian operator. Let $\phi \in L^\infty(\Omega)$ and denote by $\lambda_1^\Omega(\phi)$ the first eigenvalue of the problem

$$-\Delta u + \phi u = \lambda u, \quad u|_{\partial \Omega} = 0.$$

Clearly, $\lambda_1^\Omega(0) = \lambda_1^\Omega$. It is well-known that $\lambda_1^\Omega(\phi_n) \to \lambda_1^\Omega(\phi)$ whenever $\phi_n \to \phi$ in $L^\infty(\Omega)$, and when $\phi \leq \psi$ but $\phi \neq \psi$ in $\Omega$, then $\lambda_1^\Omega(\phi) < \lambda_1^\Omega(\psi)$.

It follows easily that, when $b(x) \geq \delta > 0$ on $\Omega$, then $\lambda(\alpha) := \lambda_1^\Omega(ab)$ is a strictly increasing function with $\lambda(0) = \lambda_1^\Omega$ and $\lambda(\alpha) \to \infty$ as $\alpha \to \infty$. Therefore, for any given $a > \lambda_1^\Omega$, there is a unique $\alpha > 0$ such that

$$a = \lambda_1^\Omega(ab).$$

We denote by $U_\alpha$ the corresponding positive normalized eigenfunction:

$$-\Delta U_\alpha + \alpha bU_\alpha = \alpha U_\alpha, \quad U_\alpha > 0, \quad U_\alpha|_{\partial \Omega} = 0, \quad \|U_\alpha\|_\infty = 1.$$  

Here and in what follows, we use the notation $\| \cdot \|_\infty = \| \cdot \|_{L^\infty(\Omega)}$.

When $b^{-1}(0) = \Omega_0$ is not empty, we assume as before that $\Omega_0 \subset \subset \Omega$ is open, connected and with smooth boundary. Then $\lambda(\alpha) = \lambda_1^\Omega(ab)$ is still strictly increasing with $\lambda(0) = \lambda_1^\Omega$, but (see [D] and [FKLM])

$$\lim_{\alpha \to \infty} \lambda(\alpha) = \lambda_1^{\Omega_0}.$$

Thus for any given $a \in (\lambda_1^\Omega, \lambda_1^{\Omega_0})$, there is a unique $\alpha > 0$ satisfying (1.2) which determines a unique $U_\alpha$ through (1.3).

We are now ready to state our main results.

**Theorem 1.1.** Suppose that $b(x) > 0$ on $\Omega$ and $a > \lambda_1^\Omega$. Let $u_p$ denote the unique positive solution of (1.1). Then the following results hold:
When \( a < \lambda_1^\Omega(b) \), we have \( u_p \to 0 \) uniformly on \( \Omega \) as \( p \to 1 + 0 \).

Moreover, as \( p \to 1 + 0 \),

\[
(p-1) \ln \|u_p\|_\infty \to \ln \alpha, \quad u_p/\|u_p\|_\infty \to U_\alpha \text{ in } C^1(\Omega),
\]

where \( \alpha \) and \( U_\alpha \) are determined by (1.2) and (1.3), respectively.

When \( a > \lambda_1^\Omega(b) \), we have \( u_p \to \infty \) uniformly on any compact subset of \( \Omega \) as \( p \to 1 + 0 \). Moreover, (1.4) holds.

(iii) When \( a = \lambda_1^\Omega(b, \Omega) \), we have \( u_p \to cU_1 \) in \( C^1(\Omega) \) as \( p \to 1 + 0 \), where \( U_1 \) is given by (1.3) with \( \alpha = 1 \) and

\[
c = \exp \left( \int bU_1^2 \ln U_1 dx / \int bU_1^2 dx \right).
\]

To understand the case that \( p \to \infty \), we need the following free boundary problem:

\[
-\Delta w = \lambda \chi_{\{w<1\}} w, \quad w > 0, \quad w|_{\partial \Omega} = 0, \quad \|w\|_\infty = 1,
\]

which also arises as a limiting problem for the degenerate predator-prey model (see [DD2]). The following result has been proved in [DD2]:

**Proposition 1.2.** For any \( a \geq \lambda_1^\Omega \), (1.5) has a unique weak solution, and when \( a < \lambda_1^\Omega \), (1.5) has no solution.

With the help of Proposition 1.2, we will prove the following:

**Theorem 1.3.** Suppose that \( b(x) > 0 \) on \( \Omega \) and \( a > \lambda_1^\Omega \). Let \( u_p \) denote the unique positive solution of (1.1). Then \( u_p \to v \) in \( C^1(\Omega) \) as \( p \to \infty \), where \( v \) is the unique positive weak solution of (1.5).

When \( \Omega_0 := b^{-1}(0) \) is a nontrivial subset of \( \Omega \), it turns out that the techniques in proving Theorems 1.1 and 1.3 are not enough. One new ingredient for dealing with this case is the following result obtained in [DD1, Lemma 2.2]:

**Lemma 1.4.** Suppose that \( \{u_n\} \subset C^1(\overline{\Omega}) \) satisfies (in the weak sense) for some positive constant \( \lambda \),

\[
-\Delta u_n \leq \lambda u_n, \quad u_n \geq 0 \text{ in } \Omega; \quad u_n|_{\partial \Omega} = 0, \quad \|u_n\|_\infty = 1.
\]

Then it has a subsequence converging weakly in \( H_0^1(\Omega) \) and strongly in \( L^q(\Omega) \) for any \( q \geq 1 \), to some \( u \) with \( \|u\|_\infty = 1 \).

**Theorem 1.5.** Suppose that \( \overline{\Omega_0} = b^{-1}(0) \) has nonempty interior which is connected with smooth boundary and \( \Omega_0 \subset \subset \Omega \). Let \( a \in (\lambda_1^\Omega, \lambda_1^{\Omega_0}) \) and denote by \( u_p \) the unique positive solution of (1.1). Then the conclusions (i)-(iii) in Theorem 1.1 hold.

As Theorem 1.5 concludes that when \( b^{-1}(0) \neq \emptyset \) and \( p \to 1 \), the behavior of \( u_p \) is the same as when \( b^{-1}(0) = \emptyset \), it is tempting to think that this is also the case when \( p \to \infty \). It turns out, however, this is not true.
Theorem 1.6. Suppose that \( \Omega_0 = b^{-1}(0) \) has nonempty interior which is connected with smooth boundary and \( \Omega_0 \subset \subset \Omega \). Let \( a \in (\lambda_1^\Omega, \lambda_1^{\Omega_0}) \) and denote by \( u_p \) the unique positive solution of (1.1). Suppose \( p_n \to \infty \) and denote \( u_n = u_{p_n} \). Then, subject to a subsequence, \( u_n \to u \) in \( L^q(\Omega) \) for all \( q \geq 1 \), where \( u \in K \) is a nontrivial nonnegative solution of the following variational inequality:

\[
\int_{\Omega} \nabla u \cdot \nabla (v - u) \, dx - \int_{\Omega} au(v - u) \, dx \geq 0, \quad \forall v \in K,
\]

where \( K := \{ w \in H_0^1(\Omega) : w \leq 1 \text{ a.e. in } \Omega \setminus \Omega_0 \} \).

In a forthcoming paper (\([\text{DD3}]\)), we will show that (1.6) has a unique positive solution for \( a \in (\lambda_1^\Omega, \lambda_1^{\Omega_0}) \), and hence \( u_p \to u \) as \( p \to \infty \) in \( L^q(\Omega) \).

Remark 1.7. From our proofs, it is easy to see that our assumptions on the smoothness of \( \partial \Omega_0 \) can be considerably weakened. For example, all our main results hold if \( \Omega_0 \) only has Lipschitz boundary.

Remark 1.8. If \( b^{-1}(0) \) consists of a single point in \( \Omega \), then Theorems 1.5 and 1.6 reduce to Theorems 1.1 and 1.3, respectively. This follows easily by checking the proofs. We intend to further consider the case that \( b^{-1}(0) \) has measure zero in \([\text{DD3}]\).

The rest of the paper consists of the proofs of our results given above. Theorem 1.1 is proved in Section 2; Theorems 1.3 and 1.5 are proved in Sections 3 and 4, respectively; Section 5 gives the proof of Theorem 1.6. The main techniques involved are various elliptic estimates and comparison principles. Several results and techniques from \([\text{DD2}]\) will be used, including fine properties of functions in Sobolev spaces and the use of variational inequalities.
2. Proof of Theorem 1.1.

Set \( M_p = \|u_p\|_\infty = \max_{\Omega} u_p \). Then it is clear that the maximum is achieved in the interior of the domain \( \Omega \), say at \( x_p \in \Omega \). Using the equation for \( u_p \) at the maximum point \( x = x_p \) we have

\[
aM_p - b(x_p)M_p^p \geq 0.
\]

Hence,

\[
M_p^{p-1} \leq a/\min_{\Omega} b.
\]

To understand the asymptotic behaviour of \( u_p \) as \( p \to 1 + 0 \), we choose an arbitrary sequence \( p_n \to 1 + 0 \) and use the notation

\[
(2.2) \quad u_n = u_{p_n}, \quad M_n = M_{p_n}, \quad \alpha_n = M_{p_n}^{p_n-1}, \quad w_n = u_n/M_n.
\]

Clearly \( w_n \) satisfies

\[
(2.3) \quad -\Delta w_n = aw_n - \alpha_n bw_n^{p_n}, \quad w_n|_{\partial \Omega} = 0.
\]

From (2.1) one sees that the right-hand side of (2.3) has a bound in \( L^\infty(\Omega) \) which is independent of \( n \). Thus, by standard elliptic estimates, \( \{w_n\} \) is bounded in \( W^{2,q}(\Omega) \) for any \( q > 1 \). By the Sobolev imbedding theorem, this implies that this sequence is compact in \( C^1(\Omega) \). Therefore, subject to a subsequence, \( w_n \to w \) in \( C^1(\Omega) \). We may also assume that \( \alpha_n \to \alpha \). Then from (2.3) we obtain, in the weak sense,

\[
-\Delta w = (a - \alpha b)w, \quad w|_{\partial \Omega} = 0.
\]

As \( w \) is nonnegative with \( \|w\|_\infty = 1 \), we necessarily have \( a = \lambda_1^\Omega(\alpha b) \) and hence \( \alpha \) is uniquely determined by (1.2) and \( w = U_\alpha \) given by (1.3). This implies that \( \alpha_n \to \alpha \) and \( w_n \to U_\alpha \) hold for the entire original sequences. Therefore, we have proved that \( M_p^{p-1} \to \alpha \) and \( u_p/M_p \to U_\alpha \) in \( C^1(\Omega) \) as \( p \to 1 + 0 \). This shows the validity of (1.4).

When \( a < \lambda_1^\Omega(b) \), we must have \( \alpha \in (0, 1) \) and it follows from

\[
(2.4) \quad \lim_{p \to 1+0} (p - 1) \ln M_p = \ln \alpha
\]

that \( M_p \to 0 \) as \( p \to 1 + 0 \). This proves Part (i) of Theorem 1.1.

When \( a > \lambda_1^\Omega(b) \), we must have \( \alpha > 1 \) and it follows from (2.4) that \( M_p \to \infty \) as \( p \to 1 + 0 \). To prove Part (ii) of Theorem 1.1, it remains to show that as \( p \to 1 + 0 \), \( u_p(x) \to \infty \) uniformly on any compact subset of \( \Omega \). To this end, for any given large number \( \beta \), we define \( V = \beta U_\alpha \) and obtain

\[
\Delta V + aV - bV^p = b(\alpha V - V^p).
\]

For those \( x \) where \( V(x) \leq 1 \), \( \alpha V - V^p \geq (\alpha - 1)V \geq 0 \); on the set \( \{ x \in \Omega : V(x) \geq 1 \} \), since \( V^p \to V \) uniformly as \( p \to 1 \), and since \( \alpha V - V \geq \alpha - 1 > 0 \), we can find \( \epsilon = \epsilon(\beta) > 0 \) small enough such that \( \alpha V - V^p > 0 \) for all \( p \in (1, 1 + \epsilon) \). Thus, for \( p \in (1, 1 + \epsilon) \), \( V \) is a lower solution to (1.1). As
any large positive constant is an upper solution of (1.1), its unique positive solution \( u_p \) must satisfy \( u_p \geq V = \beta U_\alpha \). This implies that as \( p \to 1 + 0 \), \( u_p \to \infty \) uniformly on any compact subset of \( \Omega \) and Part (ii) of Theorem 1.1 is proved.

We consider now the case that \( a = \lambda_1^\Omega (b) \). We have \( \alpha = 1 \) and hence cannot draw a conclusion for \( \lim_{p \to 1 + 0} M_p \) from (2.4). Denote \( w_p = u_p / M_p \). We have

\[-\Delta w_p = a w_p - b M_p^{p-1} w_p^p, \ w_p|_{\partial \Omega} = 0.\]

Multiply this equation by \( U_1 \), which is given by (1.3) with \( \alpha = 1 \), and integrate by parts. It results

\[
\int_\Omega (a - b) U_1 w_p \, dx = \int_\Omega (a w_p - b M_p^{p-1} w_p^p) U_1 \, dx.
\]

Hence

\[
\int_\Omega b (w_p - M_p^{p-1} w_p^p) U_1 \, dx = 0,
\]

and

\[
(2.5) \quad \int_\Omega \frac{M_p^{p-1} - 1}{p-1} b w_p^p U_1 \, dx = \int_\Omega \frac{1 - w_p^{p-1}}{p-1} b w_p U_1 \, dx.
\]

Since \( w_p \to U_1 \) as \( p \to 1 + 0 \) in \( C^1(\Omega) \), and by the Hopf boundary lemma, \( \partial U_1 / \partial \nu < 0 \) on \( \partial \Omega \), we obtain \( w_p / U_1 \to 1 \) uniformly on \( \overline{\Omega} \). It follows that

\[ ||\ln w_p - \ln U_1||_{L^\infty(\Omega)} = o(1) \]

as \( p \to 1 + 0 \). Therefore,

\[
\frac{1 - w_p^{p-1}}{p-1} w_p = 1 - e^{(p-1)(\ln U_1 + o(1))} \frac{w_p}{p-1} \to U_1 \ln U_1
\]

uniformly on \( \overline{\Omega} \) as \( p \to 1 + 0 \). From this, we see immediately that the right-hand side of (2.5) converges to

\[
\int_\Omega b U_1^2 \ln U_1 \, dx.
\]

Thus,

\[
\lim_{p \to 1+0} \int_\Omega \frac{M_p^{p-1} - 1}{p-1} b w_p^p U_1 \, dx = \int_\Omega b U_1^2 \ln U_1 \, dx,
\]

and

\[
(2.6) \quad \lim_{p \to 1+0} \frac{M_p^{p-1} - 1}{p-1} = \int_\Omega b U_1^2 \ln U_1 \, dx / \int_\Omega b U_1^2 \, dx.
\]

We show next that

\[ c := \lim_{p \to 1+0} M_p \]
exists and is uniquely determined by
\[ \ln c = \int_{\Omega} bU_1^2 \ln U_1 \, dx / \int_{\Omega} bU_1^2 \, dx. \]

We first claim that
\[ M^* := \lim_{p \to 1+0} M_p > 0, \quad M^* := \lim_{p \to 1+0} M_p < \infty. \]
Otherwise, we can find a sequence \( p_n \to 1 + 0 \) such that \( M_n := M_{p_n} \to 0 \) or \( M_n \to \infty \). In the former case, we deduce, for all large \( n \),
\[ \frac{M_n^{p_n-1} - 1}{p_n - 1} \leq \frac{\epsilon^{p_n-1} - 1}{p_n - 1} \to \ln \epsilon \]
as \( n \to \infty \), for any given \( \epsilon > 0 \). This leads to a contradiction to (2.6). In the latter case, we obtain, for all large \( n \),
\[ \frac{M_n^{p_n-1} - 1}{p_n - 1} \geq \frac{M^{p_n-1} - 1}{p_n - 1} \to \ln M \]
as \( n \to \infty \), for any given \( M > 0 \). This also leads to a contradiction to (2.6). Thus, \( 0 < M_* \leq M^* < \infty \). For any given small \( \epsilon > 0 \), a similar argument to the above leads to
\[ \ln(M_* + \epsilon) \geq \int_{\Omega} bU_1^2 \ln U_1 \, dx / \int_{\Omega} bU_1^2 \, dx, \]
\[ \ln(M^* - \epsilon) \leq \int_{\Omega} bU_1^2 \ln U_1 \, dx / \int_{\Omega} bU_1^2 \, dx. \]
Thus we necessarily have
\[ M_* = M^* = c = \exp \left( \int_{\Omega} bU_1^2 \ln U_1 \, dx / \int_{\Omega} bU_1^2 \, dx \right), \]
and \( u_p \to cU_1 \) as \( p \to 1 + 0 \) in \( C^1(\Omega) \). This finishes the proof of Theorem 1.1.

3. Proof of Theorem 1.3.

We clearly still have (2.1). Let \( p_n \) be a sequence converging to \( \infty \) and use the notations in (2.2). We find that \( u_n \) satisfies (2.3) whose right-hand side has a bound in \( L^\infty(\Omega) \) which is independent of \( n \). Thus, as in Section 2, subject to a subsequence, \( w_n \to w \) in \( C^1(\Omega) \).

The equation satisfied by \( w_n \) can also be written as
\[ -\Delta w_n = a w_n - b u_n^{p_n-1} w_n, \quad w_n|_{\partial \Omega} = 0. \]
From (2.1) we deduce
\[ 0 \leq u_n^{p_n-1} \leq a / \min_{\Omega} b. \]
Hence, by passing to a subsequence, we may assume that $bu^{p_n-1}_n \to \psi$ weakly in $L^2(\Omega)$. Clearly we must have $0 \leq \psi \leq \|b\|_\infty a/\min_{\Omega} b$. Passing to the weak limit in (3.1) we find that $w$ is a nontrivial weak solution to the problem (3.2)

$$-\Delta w = (a - \psi)w, \ w|_{\partial \Omega} = 0.$$  

As $a - \psi \in L^\infty(\Omega)$, it follows from the Harnack inequality that $w(x) > 0$ in $\Omega$.

From (2.1) we obtain

$$M_n \leq \left(\frac{a}{\min b}\right)^{1/(p_n-1)} \to 1 \text{ as } n \to \infty.$$  

It follows that $\lim_{n \to \infty} M_n \leq 1$. If $\lim_{n \to \infty} M_n < 1$, then by passing to a subsequence, we may assume that $M_n \leq 1 - \epsilon$ for all $n$ and some $\epsilon > 0$. It follows then $u_n^{p_n-1} \leq (1 - \epsilon)^{p_n-1} \to 0$ as $n \to \infty$. Hence $\psi = 0$ and $w$ is a positive solution to $-\Delta w = aw, \ w|_{\partial \Omega} = 0$. This implies that $a = \lambda_1^\Omega$, contradicting our assumption that $a > \lambda_1^\Omega$. Thus we have proved that $M_n \to 1$ as $n \to \infty$. It follows that $u_n \to w$ in $C^1(\Omega)$.

Let $\Omega_1 := \{x \in \Omega : w(x) < 1\}$. Then for any $x \in \Omega_1$, we can find $\delta > 0$ such that $u_n(x) < 1 - \delta$ for all large $n$. It follows that $0 \leq u_n(x)^{p_n-1} \leq (1 - \delta)^{p_n-1} \to 0$ as $n \to \infty$. Thus we must have $\psi = 0$ a.e. on $\Omega_1$. On the rest of $\Omega$, $w = 1$ and we necessarily have $\Delta w = 0$. (Here we regard $w$ as a member of $W^{2,q}(\Omega), q > 1$.) Thus from (3.2), we deduce $\psi = a$ a.e. on $\Omega \setminus \Omega_1$. Therefore, $w$ satisfies

$$(3.3) \quad -\Delta w = a \chi_{\{w<1\}} w, \ w > 0, \ w|_{\partial \Omega} = 0, \ |w|_\infty = 1.$$  

By Proposition 1.2, problem (3.3) has a unique solution $v$. Hence $u_n \to v$ in $C^1(\Omega)$ for the entire original sequence. This implies that $u_p \to v$ in $C^1(\Omega)$ as $p \to \infty$. The proof of Theorem 1.3 is complete.

4. **Proof of Theorem 1.5.**

We will mainly follow the lines of the proof of Theorem 1.1. The main difficulty is that the estimate (2.1) is of no use anymore and therefore it is unclear whether $\{\alpha_n\}$ is still bounded. We will use Lemma 1.4 to overcome this difficulty.

Let $p_n$ be an arbitrary sequence of numbers converging to $1+0$. We employ the notations in (2.2) and find that $w_n$ meets the conditions in Lemma 1.4. Hence, by passing to a subsequence, we may assume that $w_n \to w$ weakly in $H_0^1(\Omega)$, strongly in $L^2(\Omega)$ for any $q \geq 1$, and $|w|_\infty = 1$.

We claim that $\{\alpha_n\}$ is bounded. Otherwise, by passing to a subsequence, we may assume that $\alpha_n \to \infty$. Now we multiply (2.3), the equation satisfied by $w_n$, by $\phi/\alpha_n$ with $\phi \in C_0^\infty(\Omega)$ and integrate by parts. We obtain

$$(\alpha_n)^{-1} \int_\Omega w_n(-\Delta \phi)dx = (\alpha_n)^{-1} \int_\Omega aw_n \phi dx - \int_\Omega bw_n^{p_n} \phi dx.$$  

Letting \( n \to \infty \), we deduce
\[
\int_{\Omega} bw \phi \, dx = 0.
\]
As \( \phi \) is arbitrary, this implies that \( bw = 0 \) in \( \Omega \). Hence, \( w = 0 \) on \( \Omega \setminus \Omega_0 \). Since \( w \in H^1_0(\Omega) \) and \( \partial \Omega_0 \) is smooth, this implies that \( w|_{\Omega_0} \in H^1_0(\Omega_0) \).

Multiplying the equation for \( w_n \) by an arbitrary \( \phi \in C^\infty_0(\Omega_0) \) and integrating by parts, we obtain
\[
\int_{\Omega_0} \nabla w_n \cdot \nabla \phi \, dx = \int_{\Omega_0} a w_n \phi \, dx.
\]
Passing to \( n \to \infty \) we obtain
\[
\int_{\Omega_0} \nabla w \cdot \nabla \phi \, dx = \int_{\Omega_0} a w \phi \, dx.
\]
Thus \( w|_{\Omega_0} \) is a weak solution of the problem
\[
-\Delta u = au, \quad u|_{\partial \Omega_0} = 0.
\]
As \( w = 0 \) on \( \Omega \setminus \Omega_0 \) and \( \|w\|_\infty = 1 \), \( w|_{\Omega_0} \) is nonnegative and not identically zero. Hence we must have \( a = \lambda_1^{\Omega_0} \), contradicting our assumption that \( a < \lambda_1^{\Omega_0} \). This proves our claim that \( \{\alpha_n\} \) is bounded.

The rest of the proof follows that of Theorem 1.1 except that to prove \( u_p \geq \beta U_\alpha \), we use Lemma 2.1 of [DM] (which holds for \( C^1 \) functions).

5. Proof of Theorem 1.6.

It turns out that Lemma 1.4 is not enough for our proof of Theorem 1.6. We will need some fine properties of the limiting function \( u \) in Lemma 1.4 and of functions in \( H^1(\mathbb{R}^N) \). These fine properties have already been used in [DD2] and we simply collect them in the following lemma:

**Lemma 5.1.** Let \( u \) and \( u_n \) be as in Lemma 1.4. Then the following conclusions hold:

(i) \( \tilde{u}(x) = \lim_{r \to 0} \int_{B_r(x)} u(y) \, dy / |B_r(x)| \) exists for each \( x \in \Omega \), where \( B_r(x) \) denotes the ball with center \( x \) and radius \( r \), and \( |B_r(x)| \) stands for the volume of \( B_r(x) \). Moreover, \( u = \tilde{u} \) a.e. in \( \Omega \).

(ii) \( \tilde{u} \) is upper semi-continuous (u.s.c. for short) on \( \Omega \), and for each \( x_0 \in \Omega \) and any given \( \epsilon > 0 \), we can find a small ball \( B_r(x_0) \subset \Omega \) such that for all large \( n \),
\[
u_n(x) \leq \tilde{u}(x_0) + \epsilon, \quad \forall x \in B_r(x_0).
\]

(iii) If \( v \in H^1(\mathbb{R}^N) \), then \( \tilde{v}(x) = \lim_{r \to 0} \int_{B_r(x)} v(y) \, dy / |B_r(x)| \) exists for all \( x \in \mathbb{R}^N \) except possibly for a set of \((1,2)\)-capacity 0. Moreover, \( \tilde{v} = v \) a.e. in \( \mathbb{R}^N \) and if \( \tilde{v} \) vanishes on a closed set \( A \) in \( \mathbb{R}^N \) (except for a subset of \( A \) of capacity zero), then there exists a sequence of functions...
\( \phi_n \in H^1(R^N) \) such that each \( \phi_n \) vanishes in a neighbourhood of \( A \) and \( \phi_n \to \tilde{v} \) in \( H^1(R^N) \).

Let us now come back to the proof of Theorem 1.6. Let \( p_n \) be a sequence converging to \( \infty \) and use the notations in (2.2). Then as before, by Lemma 1.4, subject to a subsequence, \( w_n \to w \) weakly in \( H^1_0(\Omega) \) and strongly in \( L^q(\Omega) \) for any \( q \geq 1 \), and \( \|w\|_\infty = 1 \).

Claim 1. \( \{M_n\} \) is bounded.

Proof. Since \( a < \lambda_1^{\Omega_0} \), we can find a small \( \delta \)-neighborhood \( \Omega_\delta \) of \( \Omega_0 \) such that \( a < \lambda_1^{\Omega_\delta} \). Let \( \phi_\delta \) denote the normalized positive eigenfunction corresponding to \( \lambda_1^{\Omega_\delta} \):

\[-\Delta \phi_\delta = \lambda_1^{\Omega_\delta} \phi_\delta, \quad \phi_\delta|_{\partial \Omega} = 0, \quad \|\phi_\delta\|_\infty = 1,\]

and let \( \psi \in C^2(\overline{\Omega}) \) be an extension of \( \phi_\delta|_{\Omega_\delta/2} \) to \( \overline{\Omega} \) such that \( \eta := \min_{\Omega} \psi > 0 \).

We find, for any positive constant \( Q \),

\[-\Delta (Q\psi) + a(Q\psi) - b(Q\psi)^p \leq (a - \lambda_1^{\Omega_\delta})Q\psi < 0, \quad \forall x \in \Omega_\delta/2,\]
\[-\Delta (Q\psi) + a(Q\psi) - b(Q\psi)^p = Q(\Delta \psi + a\psi) - bQ^p\psi^p, \quad \forall x \in \Omega \setminus \Omega_\delta/2.\]

Let \( \xi = \inf_{\Omega \setminus \Omega_\delta/2} b \) and

\[Q_p := \left[ \xi^{-1} \sup_{\Omega} (\Delta \psi + a\psi)\eta^{-p} \right]^{1/(p-1)}.\]

We easily see that for \( Q = Q_p \),

\[-\Delta (Q\psi) + a(Q\psi) - b(Q\psi)^p \leq 0, \quad \forall x \in \Omega.\]

Therefore \( Q_p \psi \) is an upper solution of (1.1). As (1.1) has arbitrarily small positive lower solutions, its unique positive solution \( u_p \) must satisfy \( u_p \leq Q_p \psi \). Clearly \( Q_p \to 1/\eta \) as \( p \to \infty \). Thus, for any \( p_0 > 1 \), \( \{M_p : p \geq p_0 \} \) is bounded. In particular, \( \{M_n\} \) is bounded. This proves Claim 1.

By passing to a subsequence, we may assume that \( M_n \to c \in [0, \infty) \) as \( n \to \infty \).

Claim 2. \( c \geq 1 \).

Proof. Let \( v_n \) be the unique solution of

\[-\Delta v = av - \|b\|_\infty v^{p_n}, \quad v|_{\partial \Omega} = 0.\]

By Theorem 1.3 we know \( \|v_n\|_\infty \to 1 \). On the other hand, a simple comparison argument shows \( u_n \geq v_n \). Hence \( c \geq 1 \).

Claim 3. \( w \leq 1/c \) a.e. in \( \Omega \setminus \Omega_0 \).
Proof. Otherwise the set \( \{ x \in \Omega \setminus \Omega_0 : w(x) > 1/c \} \) has positive measure and we can find some \( c_1 > 1/c \) such that \( \Omega_1 := \{ x \in \Omega \setminus \Omega_0 : w(x) \geq c_1 \} \) has positive measure. As \( w_n \rightarrow w \) in \( L^2(\Omega) \), by passing to a subsequence, \( w_n \rightarrow w \) a.e. in \( \Omega \). Hence, by Egorov’s theorem, we can find a subset of \( \Omega_1 \), say \( \Omega_2 \) which has positive measure and such that \( w_n \rightarrow w \) uniformly on \( \Omega_2 \). It follows that \( u_n \rightarrow cw \) uniformly on \( \Omega_2 \). Thus, there exists \( \epsilon > 0 \) such that for all large \( n \), \( u_n \geq 1 + \epsilon \) on \( \Omega_2 \).

Let \( \phi \in C_0^\infty(\Omega) \) be an arbitrary nonnegative function, and multiply the equation for \( w_n \) by \( \phi \) and integrate over \( \Omega \). It results
\[
\int_\Omega w_n(-\Delta \phi) = a \int_\Omega w_n \phi - \int_\Omega bu_n^{p_n-1}w_n \phi.
\]
Hence, for all large \( n \),
\[
(1 + \epsilon)^{p_n-1} \int_{\Omega_2} bw_n \phi \leq \int_{\Omega_2} bu_n^{p_n-1}w_n \phi \leq \int_\Omega w_n(\Delta \phi) + a \int_\Omega w_n \phi.
\]
Dividing the above inequality by \((1 + \epsilon)^{p_n-1}\) and letting \( n \rightarrow \infty \), we deduce
\[
\int_{\Omega_2} bw \phi = 0.
\]
It follows that \( w = 0 \) a.e. in \( \Omega_2 \), contradicting the assumption that \( w \geq c_1 \) there. This proves Claim 3.

Using \( u_n = M_n w_n \) and denoting \( \hat{u} = cw \), we see from Lemma 5.1 and Claims 1-3 above that the following result holds:

**Lemma 5.2.**

(i) \( \{\|u_n\|_\infty\} \) is bounded.
(ii) Subject to a subsequence, \( u_n \rightarrow \hat{u} \) weakly in \( H_0^1(\Omega) \) and strongly in \( L^q(\Omega) \), \( \forall q \geq 1 \).
(iii) \( \hat{u} \leq 1 \) a.e. in \( \Omega \setminus \Omega_0 \) and \( \|\hat{u}\|_\infty \geq 1 \).
(iv) \( \bar{u}(x) := \lim_{r \rightarrow 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} \hat{u}(y)dy \) exists for every \( x \in \Omega \).
(v) \( \bar{u}(x) \) is u.s.c. on \( \Omega \) and \( \bar{u} = \hat{u} \) a.e. in \( \Omega \).
(vi) For each \( x_0 \in \Omega \) and any given \( \epsilon > 0 \), we can find a small ball \( B_r(x_0) \subset \Omega \) such that for all large \( n \),
\[
u_n(x) \leq \bar{u}(x_0) + \epsilon, \ \forall x \in B_r(x_0).
\]

We are now ready to complete the proof of Theorem 1.6. Multiplying the equation for \( u_n \) by \( \phi \in C_0^\infty(\Omega) \), we deduce
\[
\int_\Omega \nabla u_n \cdot \nabla \phi dx = a \int_\Omega u_n \phi dx - \int_\Omega b(x)u_n^{p_n} \phi dx.
\]
It follows that, subject to a subsequence,
\[
\lim_{n \rightarrow \infty} \int_\Omega b(x)u_n^{p_n} \phi dx = - \int_\Omega \hat{u} \cdot \nabla \phi dx + a \int_\Omega \hat{u} \phi dx, \ \forall \phi \in C_0^\infty(\Omega).
\]
Clearly the right-hand side of (5.1) defines a continuous linear functional on $H^1(\Omega)$:

$$T(\phi) = -\int_{\Omega} \nabla \hat{u} \cdot \nabla \phi dx + a \int_{\Omega} \hat{u} \phi dx.$$ 

Using the left-hand side of (5.1), and noticing that $b = 0$ on $\Omega_0$, we see that $T(\phi) \geq 0$ whenever $\phi \in C_0^\infty(\Omega)$ satisfies $\phi \geq 0$ on $\Omega \setminus \Omega_0$. Moreover, if $\text{supp}(\phi) \subset \{\tilde{u} < 1\} \cup \Omega_0$, where $\{\tilde{u} < 1\} := \{x \in \overline{\Omega} : \tilde{u}(x) < 1\}$, then by Lemma 5.2 (vi) and the fact that $\{\tilde{u} < 1\}$ is relatively open (due to $\tilde{u}$ being u.s.c.), we can find $\delta > 0$ such that $u_n(x) \leq 1 - \delta$ on the compact set $\text{supp}(\phi) \setminus \Omega_0 \subset \text{supp}(\phi) \cap \{\tilde{u} < 1\} \subset \Omega$ for all large $n$. Therefore, since $b = 0$ on $\Omega_0$

$$0 \leq \int_{\Omega} b(x) u_n^{p_n} \phi dx \leq \int_{\text{supp}(\phi) \setminus \Omega_0} b(x)(1 - \delta)^{p_n} \phi dx \to 0.$$ 

It follows that $T(\phi) = 0$ if $\text{supp}(\phi) \subset \{\tilde{u} < 1\} \cap \Omega_0$. Using the continuity of $T$ on $H^1(\Omega)$ and the fact that functions in $H^1_0(\Omega)$ can be approximated in the $H^1(\Omega)$ norm by functions in $C_0^\infty(\Omega)$, we find that

$$T(\phi) \geq 0, \forall \phi \in H^1_0(\Omega) \text{ satisfying } \phi \geq 0 \text{ a.e. on } \Omega \setminus \Omega_0,$$

(5.2)

$$T(\phi) = 0, \forall \phi \in H^1_0(\Omega) \text{ satisfying } \text{supp}(\phi) \subset \Omega_0 \cup \{\tilde{u} < 1\}.$$ 

By Lemma 5.2 (iii), we easily see that $\tilde{u} \leq 1$ on the open set $\Omega \setminus \overline{\Omega}_0$. We show next that $\tilde{u}$ is close to 0 near $\partial \Omega$ and $\tilde{u} \leq 1$ on $\partial \Omega_0$. By Lemma 5.2 (i), we can find $M > 0$ such that $au_n < M$ on $\Omega$ for all $n \geq 1$. Therefore

$$-\Delta u_n = au_n - b(x) u_n^{p_n} \leq M \text{ on } \Omega.$$ 

If $V$ is given by

$$-\Delta V = M \text{ in } \Omega, \ V|_{\partial \Omega} = 0,$$

we obtain by the maximum principle that $u_n \leq V$. It follows that $\tilde{u} \leq V$. Therefore $\tilde{u}$ is close to 0 near $\partial \Omega$.

Since $\tilde{u} \leq 1$ on $\Omega \setminus \overline{\Omega}_0$, we must have $\tilde{u} \leq 1$ on $\partial \Omega_0$ except possibly for a set of capacity zero (see, e.g., [Z] pp. 190-191).

From the above analysis, we see that it is possible to choose $\phi \in C_0^\infty(\Omega)$ such that $0 \leq \phi \leq 1$ on $\Omega$ and $\phi = 1$ on a $\delta$-neighborhood $N_\delta$ of $\{\hat{u} = 1\}$. Let $v \in K$ be arbitrary and denote $\tilde{v} = \max\{v, \phi\}$. Clearly $0 \leq \tilde{v} - v \in H^1_0(\Omega)$. Thus, by (5.2),

$$\int_{\Omega} \nabla \hat{u} \cdot \nabla (v - \hat{u}) dx = a \int_{\Omega} \hat{u} (v - \hat{u}) dx = -T(v - \hat{u}) = T(\tilde{v} - v) + T(\hat{u} - \tilde{v}) \geq T(\hat{u} - \tilde{v}).$$

Denote $u^* = \hat{u} - \tilde{v}$. Clearly $u^* \in H^1_0(\Omega)$. Now we choose $\psi \in C_0^\infty(\Omega)$ satisfying $0 \leq \psi \leq 1$ on $\Omega$, $\psi = 0$ on $\Omega \setminus N_{(2/3)\delta}$, $\psi = 1$ on $N_{(1/3)\delta}$. Then
clearly
\[ \text{supp}((1 - \psi)u^*) \subset \overline{\Omega} \setminus N_{(1/3)\delta} \subset \{ \tilde{u} < 1 \} \cup \Omega_0. \]

Hence, by (5.3),
\[ T(u^*) = T((1 - \psi)u^*) + T(\psi u^*) = T(\psi u^*). \]

As \( \psi = 0 \) on \( \Omega \setminus N_{(2/3)\delta} \), and \( \hat{v} = \max\{v, \phi\} = 1 \) a.e. on \( N_\delta \), we find that \( \psi u^* = \psi(\tilde{u} - 1) \) a.e. on \( \Omega \). Since \( \psi(\tilde{u} - 1) \) is zero outside \( N_{(2/3)\delta} \) it can be regarded as a member of \( H^1(R^N) \). It is easily seen that the representative of \( \psi(\tilde{u} - 1) \) obtained through the limiting process in Lemma 5.1 (iii) is \( \psi(\tilde{u} - 1) \). Thus we obtain
\[ T(u^*) = T(\psi u^*) = T(\psi(\tilde{u} - 1)). \]

As \( \tilde{u} \leq 1 \) on \( \overline{\Omega} \setminus \Omega_0 \) and is u.s.c., we find that the set \( A_1 := \{ \tilde{u} = 1 \} \cap (\overline{\Omega} \setminus \Omega_0) \) is closed. Let \( A_2 := R^N \setminus N_{(2/3)\delta} \) and \( A = A_1 \cup A_2 \). We know that \( \psi(\tilde{u} - 1) \) vanishes on the closed set \( A \) (except possibly for a set of capacity zero) and so by Lemma 5.1 (iii), it can be approximated in the \( H^1(R^N) \) norm by \( \phi_n \in H^1(R^N) \) with each \( \phi_n \) vanishing in a neighbourhood of \( A \). Therefore, \( \text{supp}(\phi_n) \subset \{ \tilde{u} < 1 \} \cup \Omega_0 \), and by (5.3), \( T(\phi_n) = 0 \). It follows that
\[ T(u^*) = T(\psi(\tilde{u} - 1)) = \lim_{n \to \infty} T(\phi_n) = 0. \]

We thus obtain
\[ \int_{\Omega} \nabla \hat{u} \cdot \nabla (v - \hat{u}) dx - a \int_{\Omega} \hat{u}(v - \hat{u}) dx \geq 0, \quad \forall v \in K. \]

That is to say that \( \hat{u} \in K \) is a solution of (1.6). This finishes our proof of Theorem 1.6.

References


Received July 2, 2001. The work of Dancer was partially supported by the ARC and partially done while visiting the Newton Institute at Cambridge, UK. Part of the work of Du and Ma was carried out while visiting the University of Sydney. Ma was partially supported by a grant from the National 973 Project of China and a scientific grant of Tsinghua University at Beijing.

**School of Mathematics and Statistics**  
**University of Sydney**  
NSW 2006  
**Australia**  
**E-mail address:** normd@maths.usyd.edu.au

**School of Mathematical and Computer Sciences**  
**University of New England**  
**Armidale, NSW2351**  
**Australia**  
**E-mail address:** ydu@turing.une.edu.au

**Department of Mathematical Sciences**  
**Tsinghua University**  
**Beijing 100084**  
**China**  
**E-mail address:** lma@mail.tsinghua.edu.cn
AREA-MINIMIZING MINIMAL GRAPHS OVER NONCONVEX DOMAINS

MICHAEL DORFF, DENISE HALVERSON, AND GARY LAWLOR

Sufficient conditions for which a minimal graph over a non-convex domain is area-minimizing are presented. The conditions are shown to hold for subsurfaces of Enneper’s surface, the singly periodic Scherk surface, and the associated surfaces of the doubly periodic Scherk surface which previously were unknown to be area-minimizing. In particular these surfaces are graphs over (angularly accessible) domains which have a nice complementary set of rays. A computer assisted method for proving polynomial inequalities with rational coefficients is also presented. This method is then applied to prove more general inequalities.

1. Introduction.

In this paper, we establish conditions under which a minimal graph over a certain type of nonconvex domain is area-minimizing. In particular, we consider those domains that have a “nice complementary set of rays”. Loosely speaking, a closed domain $D$ with a piecewise smooth boundary has a nice complementary set of rays if its complement can be written as the union of non-intersecting open rays emanating from $\partial D$ that are non-tangent to $\partial D$. It is shown that the compact closure of a bounded domain $D$ with connected smooth boundary has a nice complementary set of rays if and only if $D$ is angularly accessible. Such domains are a subclass of the linearly accessible domains. Linearly accessible domains are in turn associated with the close-to-convex functions, which have been studied from a geometric function theory viewpoint both in the analytic (see [5]) and the harmonic (see [1]) cases. We will show that certain regions of Enneper’s surface, the singly periodic Scherk surface, and the associated surfaces of the doubly periodic Scherk surface are area-minimizing. Furthermore, in order to prove certain inequalities necessary for our results, a computer assisted strategy is introduced. The strategy is designed to prove inequalities involving polynomials with rational coefficients. Although the required inequalities are not initially of this form, they can be modified to obtain inequalities of this form.
2. Preliminaries.

Let $M$ be an orientable surface that arises from a differentiable mapping $X$ from a domain $\Omega \subset \mathbb{R}^2$ into $\mathbb{R}^3$, so that $X(u,v) = (x(u,v), y(u,v), z(u,v))$. Fix a point $P$ on $M$. Let $t$ denote a vector tangent to $M$ at $P$ and $n$ the unit normal vector to $M$ at $P$. Then $t$ and $n$ determine a plane that intersects $M$ in a curve $\beta$. The normal curvature $\kappa_t$ at $P$ is defined to have the same magnitude as the curvature of $\beta$ at $P$ with the sign of $\kappa_t$ chosen to be consistent with the choice of orientation of $M$. The principal curvatures, $\kappa_1$ and $\kappa_2$, of $M$ at $P$ are the maximum and minimum of the normal curvatures $\kappa_t$ as $t$ ranges over all directions in the tangent space. The mean curvature of $M$ at $P$ is the average value $H = \frac{1}{2}(\kappa_1 + \kappa_2)$.

Definition 2.1.

(a) A minimal surface in $\mathbb{R}^3$ is a regular surface for which the mean curvature is zero at every point.

(b) A surface of finite area is area-minimizing if it has the least surface area of any surface having that particular boundary. A surface of infinite area is area-minimizing if each of its compact subsurfaces is area-minimizing.

Every area-minimizing surface is a minimal surface, but the converse is not true. In fact, it is difficult to prove that a specific surface is area-minimizing. However, there is the following classical result [see [6], 5.4.18 or see [11], 6.1]:

Theorem 2.2. A minimal graph over a convex domain is area-minimizing.

It follows that any subsurface of a minimal graph over a convex domain is also area-minimizing. In contrast, minimal surfaces containing area-minimizing pieces may not be area-minimizing. For example, consider the minimal surface given by

$$X(u,v) = (u - u^3/3 + uv^2, v - v^3/3 + vu^2, u^2 - v^2)$$

known as Enneper’s surface. Let $U_r$ be the disc of radius $r$ in the $u,v$-plane centered at the origin. Let $\pi_r$ denote the projection of $X(U_r)$ to the $x,y$-plane. Then $X(U_r)$ is a graph over $\pi_r$ for $r \leq 1$. Although $\pi_1$ is not convex, the convex hull of $\pi_r$ is contained in $\pi_1$ precisely for $r \leq r_0 = \sqrt{2^2/3} - 1 \approx 0.766$ and hence any subsurface of $X(U_{r_0})$ is area-minimizing. For $r_0 < r \leq 1$, the convex hull of $\pi_r$ is no longer contained in $\pi_1$ and it has been unknown whether or not $X(U_r)$ is area-minimizing for any of these $r$ values. However, it is known for $r \leq 1$ that $X(U_r)$ has least area among topological disks with the same boundary [14]. For $1 < r < \sqrt{3}$, $X(U_r)$ still has no self intersections, but is no longer a graph and does not minimize area [12]. Oprea ([13]) shows that this follows from a theorem of Schwarz,
The image of the Gauss map then contains an entire hemisphere in its interior.

The fact that a minimal graph over a nonconvex domain need not minimize area is illustrated in [11].

The types of surfaces to which we will apply our results are those that have domains that satisfy the following definition:

**Definition 2.3.** Let $D$ be a closed region in $\mathbb{R}^2$ with piecewise smooth boundary. Suppose that $\Upsilon$ is a set of rays having the following properties:

1. $\mathcal{R} \cap D = \partial \mathcal{R}$ for every $\mathcal{R} \in \Upsilon$.
2. $\mathcal{R} \cap \mathcal{R}' \subset \partial D$ for every distinct pair of rays $\mathcal{R}, \mathcal{R}' \in \Upsilon$.
3. $D = \bigcup_{\mathcal{R} \in \Upsilon} \mathcal{R}$.
4. There is a $\delta > 0$ and a set $A \subset \partial D$ so that the one-dimensional Hausdorff measure $\mathcal{H}^1(A)$ equals 0 and for all $p \in \partial D - A$, the angle between $\partial D$ and any ray $\mathcal{R} \in \Upsilon$ emanating from $p$ is defined and is at least $\delta$.

Then $\Upsilon$ is called a **nice complementary set of rays** for $D$ in $\mathbb{R}^2$. If $\Upsilon$ satisfies (1)-(3), then $\Upsilon$ is called a **complementary set** of rays for $D$ in $\mathbb{R}^2$.

The points $p \in \partial D$ at which $\partial D$ is smooth and which have a unique $\mathcal{R}(p) \in \Upsilon$ radiating from them are called the **standard points** of $\partial D$.

Note that it follows from Conditions (1) and (3) that for every $p \in \partial D$ there is at least one $\mathcal{R} \in \Upsilon$ such that $\partial \mathcal{R} = p$. Furthermore, the nonstandard points of $\partial D$ are countable. The rays radiating from a nonstandard point on the smooth parts of $\partial D$ form an angle of positive measure.

If the non-standard points were uncountable, then $\mathbb{R}^2$ would contain an uncountable collection of pairwise disjoint open sets, which is not possible.

**Lemma 2.4.** Suppose that $D$ is a compact domain with piecewise smooth connected boundary and that $\Upsilon$ is a nice complementary set of rays for $D$. Then no ray in $\Upsilon$ is contained in the tangent cone of any point of $\partial D$.

**Proof.** Let $B$ be a ball in $\mathbb{R}^2$ that contains $D$ in its interior. Note that there is a 1-1 correspondence between the rays of $\Upsilon$ and the points of $\partial B$. Let $\psi : \partial B \to \partial D$ be the map projecting $\partial B$ onto $\partial D$ via the rays of $\Upsilon$. It can be shown from the properties of a nice complementary set of rays that $\psi$ is an order preserving surjection in the sense that whenever $a, b, c, d \in \partial D$ so that $\{a, b\}$ separates $c$ and $d$, then either $\{\psi(a), \psi(b)\} \cap \{\psi(c), \psi(d)\} \neq \emptyset$ or $\{\psi(a), \psi(b)\}$ separates $\psi(c)$ and $\psi(d)$. It follows that $\psi$ is continuous and that each point preimage is a point or a connected arc.

Let $J_1, J_2, \ldots, J_n$ be smooth arcs forming $\partial D$. Define $a_n = J_1 \cap J_n$ and $a_i = J_i \cap J_{i+1}$ for $i = 1, \ldots, n - 1$. Then the endpoints of $J_i$ are $a_{i-1}$ and $a_i$ for $i = 2, \ldots, n$ and the endpoints of $J_1$ are $a_n$ and $a_1$. If $\psi^{-1}(a_i)$ is a point, denote that point as $b_i$. If not, let $b_i$ be a point in the interior of the
segment $\psi^{-1}(a_i)$. Let $K_i$ denote the arc in $\psi^{-1}(J_i)$ with endpoints $b_{i-1}$ and $b_i$ for $i = 2, \ldots, n$ or $b_n$ and $b_1$ if $i = n$.

It suffices to show that no ray in $\Upsilon$ meeting $K_i$ is contained in the tangent cone of a point of $J_i$. Suppose to the contrary that $R \in \Upsilon$ meets $K_i$ at $q$ and is contained in the tangent cone to $J_i$ at $p = \partial R$. Clearly, $q$ can not be contained in the interior of $\psi^{-1}(p)$. Hence given any set $A \subseteq J_i$ of Hausdorff measure 0, there is a sequence of points $\{q_i\} \subseteq \psi^{-1}(K_i - (A \cup \{p\}))$ so that $q_i \to q$.

Let $N(s)$ denote the outward unit normal to $J_i$ at $s$. It follows from the smoothness of $J_i$ that $N(s)$ is continuous. Let $R(t)$ denote the unit vector in the direction of the unique ray of $\Upsilon$ which meets $K$ at $t$. Then $R(t) = \frac{t - \psi(t)}{\|t - \psi(t)\|}$ is also continuous. Therefore $N(\psi(t)) \cdot R(t)$ is continuous on $K_i$. Hence $N(\psi(q_i)) \cdot R(q_i) \to 0$. This contradicts (4) of the definition of a nice complementary set of rays. Therefore, no ray in $\Upsilon$ is contained in the tangent cone of any point of $\partial D$. □

Recall that a region is linearly accessible if its complement can be written as the union of non-crossing rays. The rays are non-crossing in the sense that they may only meet at their endpoints. Furthermore, a region is said to be angularly accessible of order $\beta$, $\beta \in [0, 1]$, if its complement can be written as the union of non-crossing rays so that each ray is the bisector of a sector of measure $(1 - \beta)\pi$ that is contained in the complement of the region. Thus any angularly accessible region is linearly accessible. A linearly accessible region that is not angularly accessible of order $\beta < 1$, is said to be strictly linearly accessible.

We will say that a closed domain $D$ is angularly (or linearly) accessible if its interior is angularly (or linearly) accessible. Given an angularly accessible domain $D$ of order $\beta$, a set of rays satisfying the definition is called an access set of rays of order $\beta$ for $D$.

Clearly the concepts of a domain being strictly angularly accessible versus having a nice complementary set of rays are related. A priori, however, it is not obvious that either condition implies the other, since:

1) A measure zero set of rays is excluded from the angle requirement in Definition 2.3.

2) The angle requirement in Definition 2.3 is local, while the requirement that a wedge miss an entire region is global.

3) Angular accessibility does not require rays to have their endpoints on the boundary of the domain; there could be “feathers” where a ray $R$ extending from $D$ is not a ray of the access set, but some of the rays of the access set extend out from $R$.

The following theorem, however, reconciles these differences for piecewise smooth domains:
Theorem 2.5. A compact domain $D$ with piecewise smooth connected boundary has a nice complementary set of rays if and only if $D$ is angularly accessible of order $\beta < 1$.

Proof. Suppose that $D$ has a nice complementary set of rays, $\Upsilon$. Clearly, $D$ is linearly accessible. Assume that $D$ is strictly linearly accessible. As in Lemma 2.4, let $B$ be a ball in $\mathbb{R}^2$ that contains $D$ and let $\psi : S \to \partial D$ be the map projecting $\partial D$ onto $\partial D$ via the rays of $\Upsilon$. As discussed in Lemma 2.4, for each $p \in \partial D$, $\psi^{-1}(p)$ is a point or a connected arc in $\partial D$.

Suppose there is a sequence \{R_i\} $\subset \Upsilon$ so that the $\frac{1}{\beta}$-sector centered at $R_i$, $S(R_i, \frac{1}{\beta})$, meets $D$ at more than one point. Without loss of generality assume that $R_i \to R$. Let $p_i = \partial R_i$. Then $S(R_i, \frac{1}{\beta}) \cap \partial D - p_i \neq \emptyset$. Choose $q_i \in S(R_i, \frac{1}{\beta}) \cap \partial D - p_i$. Without loss of generality, we may assume that there is a smooth arc $J$ of $\partial D$ containing $\{p_i\}$ and $\{q_i\}$. Let $p = \partial R$. Since $J \cap R = p$, then $p_i \to p$ and $q_i \to p$. By the smoothness of $J$ there are points $r_i \in J$ between $q_i$ and $p_i$ so that the direction of the line tangent to $J$ at $r_i$ is $\frac{1}{\beta}$-close to the direction of $R_i$. Since $r_i \to p$ it follows from the smoothness of $J$ that $R$ is tangent to $J$. This is a contradiction to Lemma 2.4. Therefore $D$ is angularly accessible of order $\beta$ for some $\beta < 1$.

Conversely, suppose that $D$ is angularly accessible of order $\beta < 1$. Let $\Upsilon^*$ be an access set of rays of order $\beta$ for $D$. A nice complementary set of rays is obtained by first noting that for each point of $p \in \partial D$ there is a ray $R_p \in \Upsilon$ such that $p = \partial R_p$. Let $S_p$ be the component of the complement of $\text{int}(D) \cup (\bigcup\{R_q \mid q \neq p\})$ which contains $R_p$. Let $\Gamma_p$ be the set of all rays emanating from $p$ into $S_p$. Then the set $\Upsilon = \bigcup \Gamma_p$ is a nice complementary set of rays.

Lewandowski [10] has shown that an analytic function on the unit disc is close-to-convex if and only if its image domain is linearly accessible. There is a similar relationship between certain analytic functions and image domains that are angularly accessible of order $\beta < 1$ [8].

3. Results.

As was previously mentioned, it is well-known that a minimal surface that is a graph over a convex domain has least area among all surfaces (graphs or not) having the same boundary. In this section we provide a new area minimization result for graphs over nonconvex domains. We begin by outlining a proof by calibrations for the convex domain case (found in [11]).

Theorem 3.1. Let $D$ be a convex domain in $\mathbb{R}^2$. Let $f(x, y)$ be a function defined on $D$ whose graph $M$ is a regular minimal surface in $\mathbb{R}^3$. Then $M$ has least area among all surfaces having the same boundary as $M$.

Proof. Let $C$ be the vertical cylinder over $D$. Define a differential 2-form $\phi$ at points $(x, y, f(x, y)) \in M$ by letting it be the unit length dual of the
tangent plane, given by
\[ \phi = \frac{1}{\sqrt{1 + f_x^2 + f_y^2}} (-f_x dy dz - f_y dz dx + dx dy). \]

Extend \( \phi \) to all of \( C \) by letting it be constant in the vertical direction:
\[ \phi(x, y, z) = \phi(x, y, f(x, y)). \]
A straightforward calculation shows that a form \( \phi \) defined in this way is closed if and only if \( M \) is a minimal surface.

Now let \( S \) be any other surface whose boundary is the same as that of \( M \). If \( S \) lies entirely in \( C \), so that \( \phi \) is defined on \( S \), then we have
\[ \text{Area}(M) = \int_M \phi = \int_S \phi \leq \text{Area}(S). \]

The middle equation is by Stokes’ theorem, and the last inequality follows since \( \phi \) returns a value less than or equal to 1 when applied to tangent planes of \( S \).

If \( S \) does not lie entirely in \( C \), we employ the nearest point projection \( \Pi \) onto \( C \). Since \( \Pi \) does not increase surface area, we have
\[ \text{Area}(M) = \int_M \phi = \int_{\Pi(S)} \phi \leq \text{Area}(\Pi(S)) \leq \text{Area}(S). \]

□

Now if \( D \) is a nonconvex set, then for any projection \( \Pi \) onto \( C \), the last inequality above may not hold. However, under certain circumstances we can omit \( \text{Area}(\Pi(S)) \) from the inequality string and still obtain
\[ \text{Area}(M) = \int_M \phi = \int_{\Pi(S)} \phi \leq \text{Area}(S). \]

This is because, although \( \Pi(S) \) may have more area than \( S \), the integration of \( \phi \) may not be counting all the area of \( \Pi(S) \).

**Theorem 3.2.** Let \( D \subset \mathbb{R}^2 \) be a closed region with a nice complementary set of rays \( \Upsilon \). Let \( M \) be a minimal surface in \( \mathbb{R}^3 \) which is a graph of a function \( f(x, y) \) defined on \( D \). For every standard \( p \in \partial D \), suppose that
\[ |n(p) \cdot N(p)| \leq R(p) \cdot N(p) \]
where \( n(p) \) is the unit normal to \( M \) at \( p \), \( N(p) \) is the outward unit normal to \( \partial D \) at \( p \) naturally included into \( \mathbb{R}^2 \times \{0\} \), and \( R(p) \) is the unit normal in the direction of \( R(p) \in \Upsilon \) emanating from \( p \) also naturally included into \( \mathbb{R}^2 \times \{0\} \). Then \( M \) has least area among all surfaces (graphs or not) having the same boundary as \( M \).

**Proof.** Basic idea: Construct \( \phi \) to be the calibration for \( M \) that is dual to the tangent planes of \( M \) and extend \( \phi \) to the cylinder over \( D \) so that \( \phi \) is constant in the \( z \) direction, as before. For comparison surfaces that go outside the cylinder, project onto the cylinder in the direction defined
by the rays of $\Upsilon$. This projection stretches area (the bad news), but then integration of $\phi$ counts less than the projected area (the good news). The inequality in the hypothesis guarantees that the latter factor outweighs the former.

More formally, let $C$ be the cylinder $D \times \mathbb{R}$. Let $\Pi$ be the projection arising naturally from the choice of rays. That is, at points of $C$, $\Pi$ is the identity and at points away from $C$, $\Pi(x, y, z) = (x_0, y_0, z)$ where $(x, y) \in \mathcal{R}_\alpha$ and $(x_0, y_0) = \partial \mathcal{R}_\alpha$.

Now let $S$ be any surface (integral current) with the same boundary as $M$. Let $S_1 = S \cap C$, and $S_2 = S - S_1$. Then

$$
M(S) = M(S_1) + M(S_2)
$$

$$
\partial(\Pi \# S) = \Pi \# (\partial S) = \partial S = \partial M
$$

$$
M(M) = \int_M \phi = \int_{\Pi \# S} \phi = \int_{\Pi \# S_1} \phi + \int_{\Pi \# S_2} \phi
$$

$$
= \int_{S_1} \phi + \int_{\Pi \# S_2} \phi \leq M(S_1) + \int_{\Pi \# S_2} \phi.
$$

It remains to establish that

$$
\int_{\Pi \# S_2} \phi \leq M(S_2).
$$

Since $\Pi \# S_2$ lies on the boundary of $C$, and since $\phi$ is dual to the tangent planes of $M$, then $\phi$ applied to the tangent plane of $\Pi \# S_2$ at a point $(p, z) = (x_0, y_0, z)$ will equal $\pm n(p) \cdot N(p)$, where $n(p)$ is a unit normal to $M$ at $(p, f(p))$ and (as before) $N(p)$ is the outward unit normal to $\partial D$. Thus,

$$
\int_{\Pi \# S_2} \phi \leq \int_{\Pi \# S_2} |n(p) \cdot N(p)| \leq \int_{\Pi \# S_2} R(p) \cdot N(p).
$$

To compare the last integral with the mass of $S_2$, let $S_2$ denote the support of the current $S_2$, $\mu(x)$ its multiplicity function and $J$ the Jacobian of $\Pi$. Define $g = 1/J$ when $J \neq 0$ and $g = 0$ when $J = 0$. From the definition of a nice complementary set of rays, $R(p) \cdot N(p)$ is defined and is bounded away from 0 for almost all $p \in \partial D$. Therefore $\Pi$ is Lipschitz and it follows from the area-coarea formula that

$$
M(S_2) = \int_{S_2} \mu \geq \int_{S_2} g(x) J(x) \mu(x) dH^2
$$

$$
\geq \int_{\Pi(S_2)} \sum_{x \in \Pi^{-1}(p) \cap S_2} g(x) \mu(x) dH^2.
$$

Note that the second inequality would be an equation if we used the (potentially smaller) Jacobian of the function $\Pi$ restricted to $S_2$. 

Now if the rays $R_\alpha$ near $p \in \partial D$ are parallel, then $J = 1/(R(p) \cdot N(p))$ on those rays, as seen in Figure 1. Otherwise the rays diverge (since they can’t cross), so that $J \leq 1/(R(p) \cdot N(p))$. Thus,

\[
\int_{\Pi_{\#}S_2} R(p) \cdot N(p) = \int_{\Pi(S_2)} \mu R(p) \cdot N(p) \, d\mathcal{H}^2 \\
\leq \int_{\Pi(S_2)} \sum_{x \in \Pi^{-1}(p) \cap S_2} g(x)\mu(x) \, d\mathcal{H}^2.
\]

The last inequality may be strict for two reasons: If $R(p) \cdot N(p) < 1/J$, or if $\Pi_{\#}$ causes cancellation of sheets of $S_2$. Finally, there are only a countable number of points $p \in \partial D$ such that the Jacobian on $\Pi^{-1}(p)$ is zero, so even though $g = 0 < R(p) \cdot N(p)$ at these points, this does not affect the integral inequality. \hfill \Box

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figure1.png}
\caption{For parallel rays, $J_1 \Pi(x) = \frac{c}{a} \leq \frac{c}{b} = \sec \theta = \frac{1}{N \cdot R}$. A similar picture in $\mathbb{R}^3$ shows that $J_2 \Pi(x) = \frac{1}{N \cdot R}$.}
\end{figure}

In higher dimensions, it is not sufficient for the rays $R_\alpha$ to be non-intersecting, because that does not guarantee that the distance between them is always increasing. (They might be skew, with points of closest approach somewhere away from $C$.) One must require that the rays be either parallel or diverging from each other, and then the theorem goes through as before.

We finish this section with some results that will allow us to more easily verify that a given set of rays satisfies the conditions of Theorem 3.2 in the applications of the next section.

Let $D$ be a closed domain. A set of non-crossing rays $\mu_1, \mu_2, \ldots, \mu_k$ emanating from $\partial D$ so that $\mu_i \cap D = \partial \mu_i$ is said to partition the complement
of \( D \). The closures of the connected components of \( \mathbb{R}^2 - (D \cup \bigcup_{i=1}^{k} \mu_i) \) are called the sections of the partition.

**Definition 3.3.** Let \( S \) be a section of the complement of a closed domain \( D \) such that \( C = \partial D \cap S \) is piecewise smooth. Suppose that \( \Upsilon \) is a set of rays having the following properties:

1. \( \mathcal{R} \cap D = \partial \mathcal{R} \) for every \( \mathcal{R} \in \Upsilon \).
2. \( \mathcal{R} \cap \mathcal{R}' \subset C \) for every \( \mathcal{R}, \mathcal{R}' \in \Upsilon \).
3. \( S = \bigcup_{\mathcal{R} \in \Upsilon} \mathcal{R} \).
4. There is a \( \delta > 0 \) and a set \( A \subset \partial D \) so that the one-dimensional Hausdorff measure \( H^1(A) \) equals 0 and for all \( p \in \partial D - A \), the angle between \( C \) and any ray \( \mathcal{R} \in \Upsilon \) is defined and is at least \( \delta \).

Then \( \Upsilon \) is called a nice complementary set of rays for \( D \) in \( S \). If \( \Upsilon \) satisfies (1)-(3), then \( \Upsilon \) is called a complementary set of rays for \( D \) in \( S \).

Given a ray \( \mathcal{R} \) in \( \mathbb{R}^2 \), let \( \vartheta(\mathcal{R}) \) denote the direction of \( \mathcal{R} \) with respect to the \( x \)-axis. The first proposition is an observation.

**Proposition 3.4.** Suppose that \( D \) is a closed domain with connected boundary in \( \mathbb{R}^2 \), \( \mu_1, \mu_2, \ldots, \mu_k \) is a set of rays that partition the complement of \( D \) so that \( \vartheta(\mu_1) \leq \vartheta(\mu_2) \leq \cdots \leq \vartheta(\mu_k) \). Let \( S_i \) denote the section of the partition bounded between \( \mu_i \) and \( \mu_{(i+1 \text{ mod } k)} \). Suppose \( \Upsilon_i \) is a (nice) complementary set of rays for \( D \) in \( S_i \). Then \( \Upsilon = \bigcup \Upsilon_i \) is a (nice) complementary set of rays for \( D \) in \( \mathbb{R}^2 \). Moreover, if \( \mu_1, \mu_2, \ldots, \mu_k \) partition the complement of \( D \) into congruent sections then a given \( \Upsilon_i \) extends to a (nice) complementary set of rays for \( D \) in all of \( \mathbb{R}^2 \) via the congruence relation.

To determine if a particular set of rays in a section \( S \) of the complement of a closed domain \( D \) is a complementary set of rays for \( D \) in \( S \), we have the following definition and proposition:

**Definition 3.5.** Suppose that \( D \) is a closed domain with connected boundary in \( \mathbb{R}^2 \), \( \mu_1 \) and \( \mu_2 \) are rays of a partition of the complement of \( D \) that bounds a section \( S \), and \( \mathbf{R}_0 \) is a unit vector that is in the direction of a ray contained in \( S \). For each \( p \in \partial D \cap S \), define \( \mathcal{R}(p) \) to be a ray emanating from \( p \) in the direction of \( \mathbf{R}_0 \). Let \( W_i \) be the sector in \( S \) cut out by the angle formed by \( \mu_i \) and \( \mathcal{R}(\partial \mu_i) \). Then the \( \mathbf{R}_0 \)-set of rays for \( D \) in \( S \) is the union of the set \( \{ \mathcal{R}(p) \mid p \in \partial D \cap S \} \) together with the set of all rays emanating from \( \partial \mu_i \) contained in \( W_i \) for \( i = 0, 1 \).

**Proposition 3.6.** Let \( D \) be a closed domain with connected boundary in \( \mathbb{R}^2 \). Suppose that \( S \) is a section of a partition of the complement of \( D \) bounded by \( \mu_1 \) and \( \mu_2 \), \( C = \partial D \cap S \) is a smooth connected arc, \( \mathbf{R}_0 \) is a unit vector that is in the direction of a ray contained in \( S \) and \( \mathbf{N}(p) \cdot \mathbf{R}_0 > 0 \) for all \( p \in C - \partial C \). Then \( \Upsilon \), the \( \mathbf{R}_0 \)-set of rays for \( C \) in \( S \), is a complementary
Lemma 3.7. Suppose that $C$ is parameterized by $r : [t_1, t_2] \to C$ and $l(p)$ is the line meeting $C$ at $p = r(t_0)$ in the direction of $R_0$. Define $\delta_p(q)$ to be the distance between $q$ and $l(p)$. Then $\delta_p(r(t')) \geq \delta_p(r(t))$ whenever $t' \geq t \geq t_0$ or $t' \leq t \leq t_0$.

Proof. Suppose not. Then there is a $t^* \in (t_1, t_2)$ so that $\delta_p$ has a local maximum at $r(t^*)$. Then

$$0 = \delta_p'(r(t^*)) = \nabla \delta_p(r(t^*)) \cdot r'(t^*).$$

Thus $\nabla \delta_p(r(t^*)) \perp r'(t^*)$. In addition, $\nabla \delta_p(r(t^*)) \perp l(p)$ since the direction of greatest increase of $\delta_p$ is perpendicular to $l(p)$. Therefore $l(r(t^*))$ is parallel to $r'(t^*)$ which contradicts that $N(r(t^*)) \cdot R_0 > 0$. \hfill \Box

Lemma 3.8. The map $p \to l(p)$ is a 1-1 correspondence between the points of $C$ and the parallel lines bounded between $l(a_1)$ and $l(a_2)$.

Proof. It follows from Lemma 3.7 that no two distinct points of $C$ lie on the same line bounded between $l(a_1)$ and $l(a_2)$. Thus $C$ is contained in the region bounded between $l(a_1)$ and $l(a_2)$. By the connectedness of $C$, each of the parallel lines bounded between $l(a_1)$ and $l(a_2)$ meets $C$. Thus the map that assigns each point $p$ of $C$ to the line $l(p)$ containing $p$ and bounded between $l(a_1)$ and $l(a_2)$ is a 1-1 correspondence. \hfill \Box

For each $p \in C$, the ray $R(p)$ is the ray with endpoint $p$ that is contained in $l(p)$ and emanating from away from $D$. Hence $R(p) \cap D = \partial R(p)$. Since $D$ is a domain with connected boundary then $W_i$ can meet $D$ only at $a_i$. Thus for $R$ emanating from $a_i$ and contained in $W_i$, it is again the case that $R \cap D = \partial R$. Therefore Condition 1 is satisfied.

Let $T$ be that part of the region bounded by $l(a_1)$, $l(a_2)$ and $C$ which misses the interior of $D$. Then $S = W_1 \cup W_2 \cup T$. It should be clear from Lemma 3.8 that $S = \bigcup_{R \in \mathcal{Y}} R$ and that $R \cap R' \subset C$ for every $R, R' \in \mathcal{Y}$. Therefore, Conditions 2 and 3 are satisfied.

If $N(p) \cdot R_0 > 0$ for all $p \in C$ then Condition 4 follows from the compactness of $C$. \hfill \Box

In the applications of the next section we will identify partitions that divide a given domain $D$ into congruent pieces. Then for a specified section $S$ of such a partition, an $R_0$-set of rays for $\partial D \cap S$ in $S$ which satisfies the conditions of Proposition 3.6 is determined. Applying Proposition 3.4, this set of rays is extended to a nice complementary set of rays for $D$ in $\mathbb{R}^2$. 

set of rays for $D$ in $S$. Moreover, if $N(p) \cdot R_0 > 0$ for all $p \in C$, then $\mathcal{Y}$ is a nice complementary set of rays for $D$ in $S$.

Proof. The result will follow almost immediately from the next two lemmas.
4. Applications.

Theorem 3.2 allows us to prove that certain minimal graphs over nonconvex domains are area-minimizing. For our first application we will show that a particular portion of Enneper’s surface minimizes area. Recently, White proved that half of Enneper’s surface is area-minimizing [15], and previously, we mentioned that Enneper’s surface is area-minimizing over $\pi r_0$ for $r_0 = \sqrt{2^{2/3} - 1} \approx 0.766$. Using Theorem 3.2, we are able to establish that Enneper’s surface is area-minimizing over $\pi r_0$ for $r_0 = \frac{1}{3} \sqrt{3 + 2\sqrt{3}} \approx 0.847$ (see Figures 2 and 3).

**Theorem 4.1.** Enneper’s surface minimizes area among all surfaces spanning the curve $C = \{X(u,v) \mid u^2 + v^2 = r_0^2\}$ where $X(u,v)$ is given by Equation (1) and $r_0 = \frac{1}{3} \sqrt{3 + 2\sqrt{3}}$.

**Proof.** Using polar coordinates so that $(u,v) = (r \cos \theta, r \sin \theta)$, the first two coordinates for Enneper’s surface are given by

\[
x = r \cos \theta + r^3 \cos \theta - \frac{4}{3} r^3 \cos^3 \theta
\]

\[
y = r \sin \theta + r^3 \sin \theta - \frac{4}{3} r^3 \sin^3 \theta.
\]

Let $n$ denote the surface normal to Enneper’s surface and $N$ the outward unit normal to $\partial \pi r$ for a fixed $r$. Then

\[
N = \left(\frac{\cos \theta(1 - 3r^2 + 4r^2 \cos^2 \theta), \sin \theta(1 - 3r^2 + 4r^2 \sin^2 \theta), 0}{\sqrt{(r^2 + 1)^2 - 16r^2 \cos^2 \theta \sin^2 \theta}}\right), \quad \text{and}
\]

\[
n = \frac{1}{1 + r^2} \left(2r \cos \theta, -2r \sin \theta, r^2 - 1\right).
\]

Let $\theta_i = \frac{(2i-1)\pi}{4}$ for $i = 0, 1, 2, 3$. It is a straightforward calculation to show that $X(r \cos \theta, r \sin \theta)$ meets the ray emanating from the origin in the $\theta_i$ direction precisely when $\theta = \theta_i$ and that the distance from $X(r \cos \theta_i, r \sin \theta_i)$ to the origin increases in $r$. For a fixed $r$, let $\mu_i$ be the ray in the direction of $\theta_i$ emanating from $X(r \cos \theta_i, r \sin \theta_i)$ for $i = 0, 1, 2, 3$. Then $\{\mu_i\}$ partitions the complement of $\pi_r$ into congruent pieces. Let $S_1$ be the section of $\mathbb{R}^2$ bounded between $\mu_0$ and $\mu_1$. Note that $C_1 = S \cap \partial \pi_r = \{X(r \cos \theta, r \sin \theta) \mid -\pi/4 \leq \theta \leq \pi/4\}$ is a smooth curve.

Let $R_0 = (1, 0, 0)$ and let $\Upsilon_1$ be the $R_0$-set of rays for $\pi_r$ in $S_1$. Observe that for $p \in \partial C_1$,

\[
N(p) \cdot R_0 = \frac{\cos \theta(1 - 3r^2 + 4r^2 \cos^2 \theta)}{\sqrt{(r^2 + 1)^2 - 16r^2 \cos^2 \theta \sin^2 \theta}} \geq \frac{1 - r^2}{\sqrt{2(1 + r^2)}}.
\]
Hence, by Proposition 3.6, \( \Upsilon_1 \) is a nice complementary set of rays for \( \pi_r \) in \( S_1 \) when \( r < 1 \). By Proposition 3.4, \( \Upsilon_1 \) extends to \( \Upsilon \), a nice complementary set of rays for \( \pi_r \) in \( \mathbb{R}^2 \).

In order to apply Theorem 3.2 to \( \Upsilon \), we want to find the largest value of \( r > 0 \) for which

\[
|\mathbf{n} \cdot \mathbf{N}| \leq \mathbf{N} \cdot \mathbf{R}
\]

almost everywhere on \( \partial \pi_r \). Note that

\[
\mathbf{n} \cdot \mathbf{N} = \frac{2r(2\cos^2 \theta - 1)}{\sqrt{(r^2 + 1)^2 - 16r^2 \cos^2 \theta \sin^2 \theta}} \geq 0.
\]

Thus, we want to find the largest region \( \pi_r \) so that \( \mathbf{N} \cdot (\mathbf{R}_0 - \mathbf{n}) \geq 0 \) holds. In particular, we want \( r \) to satisfy

\[
\cos \theta (1 - 3r^2 + 4r^2 \cos^2 \theta) - 2r(\cos^2 \theta - \sin^2 \theta) \geq 0
\]

for \( \theta \in [-\pi/4, \pi/4] \). The value of \( \theta \in [-\pi/4, \pi/4] \) that minimizes the left-hand side of (3) satisfies

\[
\cos \theta = \frac{2 + \sqrt{1 + 9r^2}}{6r}.
\]

Substituting this value into (3), and solving for \( r \) we get that

\[
r_0 = \frac{1}{3} \sqrt{3 + 2\sqrt{3}} \approx 0.84748.
\]
Next, we apply Theorem 3.2 to determine portions of the Scherk singly periodic surface that are the minimal graphs over nonconvex domains which are area-minimizing. The Scherk singly periodic surface arises in minimal surface theory as the conjugate surface of the Scherk doubly periodic surface. Recall the following definitions:

**Definition 4.2.** If a minimal surface $X(u, v) = (x, y, z)$ is defined on a simply connected domain $\Omega \in \mathbb{C}$, then we define the *conjugate surface* or *adjoint surface*, $X^*(u, v) = (x^*, y^*, z^*)$ to $X(u, v)$ on $\Omega$ as the solution of the Cauchy-Riemann equations

$$X_u = X^*_v, \quad X_v = -X^*_u$$

in $\Omega$.

**Example 4.3.** The doubly periodic Scherk surface (see Figure 4) is given by

$$X(u, v) = \left( \frac{1}{2} \arg \left( \frac{z - i}{z + i} \right), \frac{1}{2} \arg \left( \frac{1 + z}{1 - z} \right), \frac{1}{2} \log \left| \frac{1 + z^2}{1 - z^2} \right| \right)$$

where $z = u + iv$. Its conjugate surface is the singly periodic Scherk surface (see Figure 5) given by

$$X^*(u, v) = \left( \frac{1}{2} \log \left| \frac{z + i}{z - i} \right|, \frac{1}{2} \log \left| \frac{1 - z}{1 + z} \right|, \frac{1}{2} \arg \left( \frac{1 + z^2}{1 - z^2} \right) \right).$$

[Figure 4. The doubly periodic Scherk surface with $r = 0.999$.]

[Figure 5. The singly periodic Scherk surface with $r = 0.999$.]

Now let $\pi^0_r$ denote the projection of that portion of the doubly periodic Scherk surface defined on $U_r$ to the $x, y$-plane. Likewise, let $\pi^*_r$ denote the projection of that portion of the singly periodic Scherk surface defined on $U_r$ to the $x, y$-plane. Since the doubly periodic Scherk surface is a graph over $\pi^0_r$ which is convex (see Figure 6), it is area-minimizing by Theorem 2.2. By a result by Krust [see [7] or [2]], the singly periodic Scherk surface is a minimal
graph over $\pi^*_1$. The region $\pi^*_1$ is nonconvex (see Figure 7) and it has been unknown whether or not it is area-minimizing over $\pi^*_r$ for $r > 1/\sqrt{2} \approx 0.707$; i.e., for values $r$ such that the convex hull of $\pi^*_r$ is contained in $\pi^*_1$. However, $\pi^*_r$ has a nice complementary set of rays for $r < 1$. Thus we can apply Theorem 3.2 to derive the following result:

**Theorem 4.4.** The singly periodic Scherk surface minimizes area among all surfaces spanning the curve $C = \{X(u, v) \mid u^2 + v^2 = r_0^2\}$ where $X(u, v)$ is given by Equation (4) and

$$r_0 = \frac{1}{2} \sqrt{-10 - 2\sqrt{17} + 8\sqrt{4 + \sqrt{17}} - 2\sqrt{102 + 26\sqrt{17} - 40\sqrt{4 + \sqrt{17}} - 8\sqrt{17}\sqrt{4 + \sqrt{17}}} \approx 0.770778. $$

**Proof.** Parametrizing the $u, v$-plane in polar coordinates so that $(u, v) = (r \cos \theta, r \sin \theta)$, we have

$$x = \frac{1}{4} \log \left( \frac{r^2 + 1 + 2r \sin \theta}{r^2 + 1 - 2r \sin \theta} \right), $$

$$y = \frac{1}{4} \log \left( \frac{r^2 + 1 - 2r \cos \theta}{r^2 + 1 + 2r \cos \theta} \right). $$

Let $n$ denote the surface normal to the singly periodic Scherk surface and $N$ the outward unit normal to $\partial \pi^*_r$ for a fixed $r$. Then

$$N = \left( \frac{\sin \theta[(1 - r^2)^2 + (2r \cos \theta)^2], - \cos \theta[(1 - r^2)^2 + (2r \sin \theta)^2], 0}{\sqrt{(1 - r^2)^4 + 16r^2 \cos^2 \theta \sin^2 \theta(1 - r^2 + r^4)}}, \right)$$

![Figure 6. Projection of one component of the doubly periodic Scherk surface with $r_1 = \frac{1}{\sqrt{2}}$, $r_2 = 0.770778$ and $r_3 = 0.999$.](image)

![Figure 7. Projection of the singly periodic Scherk surface with $r_1 = \frac{1}{\sqrt{2}}$, $r_2 = 0.770778$ and $r_3 = 0.999$.](image)
and

\[ n = \frac{-1}{1 + r^2} \begin{pmatrix} 2r \cos \theta, -2r \sin \theta, r^2 - 1 \end{pmatrix}. \]

Let \( \theta_i = \frac{i\pi}{2} \) for \( i = 0, 1, 2, 3 \). It is a straightforward calculation to show that \( X(r \cos \theta, r \sin \theta) \) meets the ray emanating from the origin in the \( \theta_i \) direction precisely when \( \theta = \theta_i + \frac{\pi}{2} \) and that the distance from \( X(r \cos \theta_i, r \sin \theta_i) \) to the origin increases in \( r \). For a fixed \( r \), let \( \mu_i \) be the ray in the direction of \( \theta_i \) emanating from \( X(r \cos(\theta_i + \frac{\pi}{2}), r \sin(\theta_i + \frac{\pi}{2})) \) for \( i = 0, 1, 2, 3 \). Then \( \{\mu_i\} \) partitions the complement of \( \pi_r \) into congruent pieces. Let \( S_1 \) be the section of \( \mathbb{R}^2 \) bounded between \( \mu_0 \) and \( \mu_1 \). Note that \( C_1 = S \cap \partial \pi_r = \{X(r \cos \theta, r \sin \theta) \mid \frac{\pi}{2} \leq \theta \leq \pi\} \) is a smooth curve.

Let \( R_0 = (1/\sqrt{2}, 1/\sqrt{2}, 0) \) and let \( \Upsilon_1 \) be the \( R_0 \) set of rays for \( \partial \pi_r \cap S_1 \). Observe that for \( p \in \partial \pi_r \cap S_1 \),

\[
N(p) \cdot R_0 = \frac{\sin(\theta)((1 - r^2)^2 + 2r \cos \theta)^2) - \cos \theta(1 - r^2)^2 + 2r \sin \theta)^2]}{\sqrt{2}(1 - r^2)^4 + 16r^2 \cos \theta \sin \theta(1 - r^2)^4}} \quad \geq \frac{1}{\sqrt{2}}.
\]

Hence by Proposition 3.6, \( \Upsilon_1 \) is a nice complementary set of rays for \( \pi_r \cap S_1 \) in \( S_1 \). By Proposition 3.4, \( \Upsilon_1 \) extends to \( \Upsilon_r \), a nice complementary set of rays for \( \pi_r \) in \( \mathbb{R}^2 \).

We want to find the largest value of \( r \) so that \( |n \cdot N| \leq N \cdot R \). Note that for \( \theta \in [\pi/2, \pi] \)

\[
n \cdot N = \frac{-4r \cos \theta \sin \theta(1 + r^4)}{(1 + r^2)(1 - r^2)^4 + 16r^2 \cos \theta \sin \theta(1 - r^2)^4}) \geq 0.
\]

Thus we want to find the largest region \( \pi_r \) so that \( N \cdot (R - n) \geq 0 \). In particular, we want to satisfy

\[
(1 + r^2)(\sin \theta - \cos \theta)((1 - r^2)^2 - 4r^2 \cos \theta \sin \theta) \quad + 4\sqrt{2}r \cos \theta \sin \theta(1 + r^4) \geq 0
\]

for \( \theta \in [\pi/2, \pi] \). Differentiating and using the identity \( \cos \theta \sin \theta = \frac{1}{2}[1 - (\cos \theta - \sin \theta)^2] \) we find that the value of \( \theta \) that minimizes the left-hand side of this inequality satisfies the equation

\[
\cos \theta - \sin \theta = \frac{-2\sqrt{2}(1 + r^4) \pm \sqrt{2r^8 + 12r^6 + 52r^4 + 12r^2 + 2}}{6r(1 + r^2)}.
\]
Substituting this value into (5) yields

\[
\frac{\sqrt{2}}{54r(1+r^2)^2}[(1+r^4)(r^8+36r^6+38r^4+36r^2+1)
\pm (r^8+6r^6+6r^2+1)^{\frac{1}{2}}] \geq 0.
\]

Hence, we need the expression inside of the square brackets to be greater than 0. This leads to finding the smallest positive root of

\[
54r^2(r^2+1)^2[r^{16}+20r^{14}-8r^{12}+28r^{10}-146r^8+28r^6-8r^4+20r^2+1].
\]

Factoring \( r^8 \) out of the expression in the square brackets and substituting

\[ t = r^2 + \frac{1}{r^2}, \]

gives the new expression in the square brackets as

\[
t^4 + 20t^3 - 12t^2 - 32t - 128,
\]

whose roots are \(-5 - \sqrt{17} \pm 4\sqrt{4 + \sqrt{17}} \) and \(-5 + \sqrt{17} \pm 4\sqrt{4 - \sqrt{17}}\). The smallest positive root of

\[
r^{16}+20r^{14}-8r^{12}+28r^{10}-146r^8+28r^6-8r^4+20r^2+1
\]

is therefore

\[
r_0 = \frac{1}{2} \sqrt{-10 - 2\sqrt{17} + 8\sqrt{4 + \sqrt{17}} - 2\sqrt{102 + 26\sqrt{17}} - 4\sqrt{4 + \sqrt{17}} - 8\sqrt{17}\sqrt{4 + \sqrt{17}}}.\]

In general, a conjugate surface is a minimal surface. Thus, we can construct a one-parameter family of minimal surfaces.

**Definition 4.5.** For \( \alpha \in \mathbb{R} \), the surfaces \( Z(u,v,\alpha) \) are called associated minimal surfaces to the surface \( X(u,v) \), where

\[
Z(u,v,\alpha) := X(u,v) \cos \alpha + X^*(u,v) \sin \alpha.
\]  

Associated surfaces share several nice properties. In the references above, Krust showed that if an embedded minimal surface can be written as a graph over a convex domain, then all associated minimal surfaces are graphs. Recently, Dorff [3] proved that these associated surfaces are graphs over close-to-convex domains.

**Example 4.6.** The associated surface of the doubly periodic Scherk surface for \( \alpha = \pi/4 \) along with its projection onto a close-to-convex domain are shown in Figures 8 and 9.
Using Theorem 3.2 and a computer assisted approach for proving inequalities, we now determine a whole class of subsurfaces of the Scherk associated surfaces that are graphs over nonconvex domains and are area-minimizing.

**Theorem 4.7.** For $0 \leq \alpha \leq \frac{\pi}{2}$, the associated Scherk surface minimizes area among all surfaces spanning the curve $C_\alpha = \{Z(u, v, \alpha) \mid u^2 + v^2 = r_0^2\}$ where $Z(u, v, \alpha)$ is given by Equation (6) and $r_0 = 0.77043$.

**Proof.** Let the $u,v$-plane be parameterized in polar coordinates so that $(u, v) = (r \cos \theta, r \sin \theta)$. Note that the normal $N_\alpha$ to $C_\alpha$ in the plane $z = 0$ is given by

$$N_\alpha = \frac{\tilde{N}_\alpha}{\|N_\alpha\|}$$

where $$\tilde{N}_\alpha = \cos \alpha N_0 + \sin \alpha N_{z}. $$

It follows that

$$\tilde{N}_\alpha = \cos \alpha \left( \frac{r(1-r^2) \cos \theta}{[(1-r^2)^2 + (2r \sin \theta)^2]^\frac{1}{2}}, \frac{r(1-r^2) \sin \theta}{[(1-r^2)^2 + (2r \cos \theta)^2]^\frac{1}{2}}, 0 \right)$$

$$+ \sin \alpha \left( \frac{r(r^2+1) \sin \theta}{[(1-r^2)^2 + (2r \sin \theta)^2]^\frac{1}{2}}, \frac{-r(r^2+1) \cos \theta}{[(1-r^2)^2 + (2r \cos \theta)^2]^\frac{1}{2}}, 0 \right),$$

and

$$n = \frac{-1}{1 + r^2} (2r \cos \theta, -2r \sin \theta, r^2 - 1).$$

To construct a nice complementary set of rays for $\pi^\alpha_r$ that satisfies the conditions of Theorem 3.2, we first identify a simple partition of the complement.
of $\pi_\alpha^r$ from which we construct a complementary set of rays for $\pi_\alpha^r$. A finite subset of these rays then gives a different partition of the complement of $\pi_\alpha^r$. Finally, we utilize the new partition to define a nice complementary set of rays for $\pi_\alpha^r$.

To obtain our first partition, note that the maximum $x$-coordinate for points of $\pi_\alpha^r$ occurs when the $y$ component of $\mathbf{N}_\alpha$ is 0. This occurs when

$$\tan \theta = \frac{1 + r^2}{1 - r^2} \tan \alpha.$$  

Let

$$\theta_i^* = \tan^{-1} \left[ \frac{1 + r^2}{1 - r^2} \tan \alpha \right] + \frac{\pi i}{2} \text{ for } i = 0, 1, 2, 3.$$  

Let $\phi_i = \frac{\pi i}{2}$ for $i = 0, 1, 2, 3$. Let $\hat{Z}$ denote the orthogonal projection of the map $Z$ to the $xy$-plane. Let $\mu_i^*$ be the ray in the $\phi_i$ direction emanating from $\hat{Z}(r \cos \theta_i^*, r \sin \theta_i^*, \alpha)$. Then clearly \{\mu_i^*\} partitions the complement of $\pi_\alpha^r$ into congruent sections. Let $S_0^*$ be the section bounded between $\mu_0^*$ and $\mu_1^*$. Note that

$$C_0^* = S_0 \cap \partial \pi_\alpha^r = \{ \hat{Z}(r \cos \theta, r \sin \theta, \alpha) \mid \theta_0^* < \theta < \theta_1^* \}$$  

is a smooth curve. Let $\mathbf{R}^* = (1, 0, 0)$. Then

$$\mathbf{N}_\alpha \cdot \mathbf{R}^* = \frac{r(1 - r^2) \cos \alpha \cos \theta + r(r^2 + 1) \sin \alpha \sin \theta}{\|\mathbf{N}_\alpha\|[(1 - r^2)^2 + (2r \sin \theta)^2]}.$$  

For $\theta_0^* \leq \theta \leq \frac{\pi}{2}$, it is clear that $\mathbf{N}_\alpha \cdot \mathbf{R}^* > 0$. For $\theta_1^* \leq \theta < \frac{\pi}{2}$ we have

$$- \cos \theta < \frac{1 + r^2}{1 - r^2} \tan \alpha \sin \theta.$$  

It follows that $\mathbf{N}_\alpha \cdot \mathbf{R}^* > 0$ for $\frac{\pi}{2} \leq \theta \leq \theta_1^*$. (In fact $\mathbf{N}_\alpha \cdot \mathbf{R}^* = 0$ at $\theta = \theta_1^*$.) Hence the $\mathbf{R}^*$-set of rays for $\pi_\alpha^r$ in $S_0^*$ extends to a complementary set of rays $\Upsilon^*$ for $\pi_\alpha^r$. Note that any finite set of rays in $\Upsilon^*$ forms a partition of the complement of $\pi_\alpha^r$

Now let

$$\omega(\alpha) = \cos^{-1} \left[ - \frac{19}{64} + \frac{7}{3} \cos \frac{\alpha}{2} - \frac{4}{3} \cos^2 \frac{a}{2} - \frac{17}{10} \sin \frac{a}{2} + \sin^2 \frac{a}{2} \right].$$  

It can be shown by the computer assisted method of the next section that for $\alpha \in \left[ 0, \frac{\pi}{2} \right],$

(7) \hspace{1cm} \theta_0^* \leq \omega(\alpha) \leq \theta_1^*.

Let $\omega_i = \omega(\alpha) + \frac{\pi i}{2}$ for $i = 0, 1, 2, 3$. Let $\mu_i$ be the ray emanating from $\hat{Z}(r \cos \omega_i, r \sin \omega_i, \alpha)$ in the $\phi_i$ direction. Then \{\mu_i\} $\subset \Upsilon^*$ and hence \{\mu_i\}
partitions the complement of $\pi_r^\alpha$. Let $S_0$ be the section of the complement bounded between $\mu_0$ and $\mu_1$. Note that

$$C_0 = S_0 \cap \partial \pi_r^\alpha = \{ Z(r \cos \theta, r \sin \theta, \alpha) \mid \omega_0 < \theta < \omega_1 \}$$

is a smooth curve. Let $R_\alpha = \left( \sin \frac{\alpha}{2}, \cos \frac{\alpha}{2}, 0 \right)$. Again, by using the computer assisted approach of the next section, it can be shown that for $\alpha \in \left[ 0, \frac{\pi}{2} \right]$, $\theta \in [\omega_0, \omega_1]$ and $r_0 = 0.77043$ that

$$N_\alpha \cdot R_\alpha - |n \cdot N_\alpha| > 0.$$  \hspace{1cm} (8)

Hence the $R_\alpha$-set of rays for $\pi_r^\alpha$ in $S_0$ extends to a nice complementary set of rays for $\pi_r^\alpha$. Moreover, the hypothesis of Theorem 3.2 is satisfied for $r_0 = 0.77043$ and all $\alpha \in \left[ 0, \frac{\pi}{2} \right]$. Therefore, for $0 \leq \alpha \leq \frac{\pi}{2}$, the associated Scherk surface minimizes area among all surfaces spanning the curve $C_\alpha = \{ Z(u, v, \alpha) \mid u^2 + v^2 = r_0^2 \}$ where $r_0 = 0.77043$. \hfill \Box

5. A computer assisted approach to proving inequalities.

We begin with a simple observation about polynomials: If a polynomial is dominated by its constant coefficient, then for small $x$ the value is always positive or always negative. More specifically:

**Lemma 5.1.** Let $P(x) = a_0 + a_1 x + \cdots + a_n x^n$. Suppose that

$$|a_0| > |a_1| + |a_2| + \cdots + |a_n|.$$  

Then $P(x)$ has the same sign as $a_0$ for all $x \in [-1, 1]$.

**Proof.** Let $x \in [-1, 1]$. Then

$$a_0 P(x) - a_0^2 = a_0(a_1 x + \cdots + a_n x^n)$$

$$\geq -|a_0||a_1 x + \cdots + a_n x^n|$$

$$\geq -|a_0|(|a_1||x| + \cdots + |a_n||x^n|)$$

$$\geq -|a_0|(|a_1| + \cdots + |a_n|).$$

Thus

$$\frac{a_0}{|a_0|} P(x) \geq |a_0| - (|a_1| + \cdots + |a_n|) > 0.$$ \hspace{1cm} \Box

This gives a method for proving a polynomial inequality for $x \in [-1, 1]$. We now extend Lemma 5.1 in order to prove polynomial inequalities for more general intervals $x \in [x_0 - t, x_0 + t]$. 

Lemma 5.2. Let $P(x) = a_0 + a_1 x + \cdots + a_n x^n$, and for some fixed $t > 0$ and $x_0 \in \mathbb{R}$ let $Q(x) = P(xt + x_0)$. Suppose that when expanded out, $Q(x) = b_0 + b_1 x + \cdots + b_n x^n$ and that $|b_0|$ dominates the sum of absolute values of the other coefficients $b_i$ as in Lemma 5.1. Then $P(x)$ has the same sign as $b_0$ for all $x \in [x_0 - t, x_0 + t]$.

Proof. By Lemma 5.1, $\frac{b_0}{|b_0|}Q(x) > 0$ for all $x \in [-1, 1]$. Set $s = xt + x_0$, so that when $x \in [-1, 1]$, $s \in [x_0 - t, x_0 + t]$. Since $Q(x) = P(xt + x_0)$, we get that $\frac{b_0}{|b_0|}Q(x) = \frac{b_0}{|b_0|}P(s) > 0$ for all $s \in [x_0 - t, x_0 + t]$, as desired. \qed

Lemma 5.2 gives a method by which a computer can prove a polynomial inequality on a compact interval $I$. Divide $I$ into small pieces. Notice that the smaller $t$ is, the more likely $b_0$ is to dominate the other coefficients of $Q(x) = P(xt + x_0)$, since the other coefficients all get multiplied by a positive power of $t$. If $P$ is strictly positive on all of $I$, then this method will work for a sufficiently fine subdivision.

Next we deal with polynomials of more than one variable.

Lemma 5.3. Let $x = (x_1, \ldots, x_k)$, and let $P(x)$ be a polynomial of degree $n$ with constant coefficient $a_0$. If $a_0$ dominates all the other coefficients of $P$ as in Lemma 5.1, then for all $x \in [0, 1] \times \cdots \times [0, 1]$, $P(x)$ has the same sign as $a_0$.

Proof. Essentially the same as Lemma 5.1. \qed

Lemma 5.4. Let $x = (x_1, \ldots, x_k)$ and let $P(x)$ be a polynomial of degree $n$. Fix a collection of “stretch values” $t_1, \ldots, t_k$ and “center values” $p_1, \ldots, p_k$. Let $Q(x_1, \ldots, x_k) = P(x_1 t_1 + p_1, \ldots, x_k t_k + p_k)$. Suppose that when $Q$ is expanded out, its constant coefficient $b_0$ dominates all its other coefficients as in Lemma 5.1. Then $P(x)$ has the same sign as $b_0$ for all $x \in [p_1 - t_1, p_1 + t_1] \times \cdots \times [p_k - t_k, p_k + t_k]$.

Proof. The proof follows that of Lemma 5.2. \qed

Now we describe general strategy for proving that a polynomial $P(x)$ is positive on all of a box $[a_1, b_1] \times \cdots \times [a_k, b_k]$. First, for each $i$ we set $p_i = \frac{1}{2}(a_i + b_i)$ and $t_i = \frac{1}{2}(b_i - a_i)$. We test $P$ by Lemma 5.3. If it fails, we divide the box in half in each of the $k$ directions, obtaining $2^k$ smaller boxes. We test each of these in turn. The ones that pass, we check off our list, and the ones that still fail, we subdivide in half again. We continue this process until all boxes are small enough to pass, which will eventually be true as long as $P$ actually is strictly positive on all of the original box.

If the coefficients of $P$ and the bounding values $\{a_i\}$ and $\{b_i\}$ of the original box are all rational, then a computer program such as Mathematica can do all of the calculations in exact rational arithmetic, thus eliminating all roundoff error.
6. Mathematica programs.

First we set up the two test functions, Test1Var and Test2Var, as well as the values and functions that will be plugged into them. Test1Var takes a polynomial in one variable with rational coefficients and rational minimum and maximum values prescribed for the domain, and proves (if it succeeds) that the polynomial is positive on the prescribed domain. Similarly, Test2Var does the same for a polynomial in two variables. The main difference in Test2Var is that the domain is usually more complicated to describe.

Line-by-line comments are given below for documenting Test1Var. A few further comments are later added for the features unique to Test2Var.

Note that the lines are numbered along the left side for explanation below. These line numbers are not part of the Mathematica program.

6.1. The single variable inequality test.

(1) Test1Var[f_, umin_, umax_] := (  
(2) blox = List[List[(umin + umax)/2, (umax - umin)/2]];  
(3) While[Length[blox] > 0,  
(4)  For[1 = 1, i < Length[blox] + 1, i++,  
(5)    uc = blox[[i, 1]]; du = blox[[i, 2]];  
(6)    If[f[uc] < 0, flag = "fail",  
(7)      If[2*f[uc] - Apply[Plus, Abs[Flatten[CoefficientList  
(8)        [f[u*du + uc], {u}]]]] > 0,  
(9)        flag = "drop",  
(10)       flag = "subdivide"]]];  
(11)   If[flag == "fail", Print["f is negative at ", uc, "]"]; i = 2;  
(12)   blox = List[];  
(13)   If[flag == "drop",  
(14)      blox = Join[Take[blox, i - 1], Drop[blox, i]]; i = i - 1];  
(15)   If[flag == "subdivide",  
(16)      hdu = du/2;  
(17)      blox = Join[Take[blox, i - 1],  
(18)        List[List[uc - hdu, hdu],  
(19)        List[uc + hdu, hdu]],  
(20)        Drop[blox, i]; i = i + 1]];  
(21)   Print["Number of blocks remaining: ", Length[blox]];  
(22)   If[Length[blox] > 0, Print["Width: ", N[2*hdu]]  
(23) ) ]

Comments:

Line 1: This program takes as input a polynomial P[u], and a closed interval domain (defined by minimum and maximum values of u) within which to test whether P is positive.
250 M. DORFF, D. HALVERSON, AND G. LAWLOR

Line 2: An individual block (i.e., subinterval) is represented as a list of two numbers, giving its center $u$ value and its radius (half its $u$-width). Line 1 sets up the initial list of “all” blocks, which consists of just one, the original domain.

Line 3: This While statement runs the main For loop over and over until the list of blocks remaining to be checked out is empty.

Line 4: This executes the For loop, which starts with the first block in the list and continues to the last one.

Line 5: Brings out the $i$th block to be tested.

Lines 6: If the function is negative at the center of the block, then set the flag to report this. Otherwise, go on to Line 7.

Lines 7-9: The command CoefficientList makes a list of the coefficients of the shifted and stretched polynomial $P[u \cdot du + uc]$. The list has sublist structuring in it, so we Flatten it into a single list, and then take the Absolute value of all the coefficients. The command “Apply[Plus,...]” adds these absolute values. The constant coefficient is $P[uc]$, which is nonnegative (because of the test in Line 8), and which is also included in the coefficient list, so that “2P[uc]-Apply[Plus,...]” works out to be the constant coefficient minus the sum of absolute values of the remaining coefficients. If this is positive, then the inequality test in Lemma 4 has worked for the current block, so the flag is set to drop the block from the list. Otherwise, go to Line 10.

Line 10: The current block must be subdivided. This line ends the testing of the current block.

Lines 11-12: In this case report test failure and set the values of $i$ and blox so that both the For and the While loops will terminate.

Lines 13-14: In this case remove the current block from the list to be tested, adjust the index $i$, and return to the For statement to test the next block.

Lines 15-20: In this case subdivide the current block into two blocks of half the size, insert them into the testing list, and adjust the index $i$ so that the reduced blocks aren’t tested until the next full pass through the list. Return to the For statement to test the next block in the list.

Lines 21-22: At the end of one full pass of the For statement through all the blocks, report how many blocks remain to be tested and what is the current width of blocks (if there are any). Return to the While statement, which will run the next pass if Lemma 5.3 (i.e., Lines 9-16) has not yet succeeded in eliminating all blocks from the list.

Line 23: This ends the While statement and the definition of Test1Var.

6.2. The two variable inequality test.

(1) $g[a] := \text{Take}[a, 2]$
(2) Test2Vars[f, umin, umax, vmin, vmax] := ( 
  (3) blox = List[
    (4) List[(umin + umax)/2, (vmin + vmax)/2, 
    (5) (umax - umin)/2, (vmax - vmin)/2]; 
  (6) While[Length[blox] > 0, 
    (7) For[i = 1, i < Length[blox] + 1, i++, 
      (8) uc = blox[[i, 1]]; vc = blox[[i, 2]]; du = blox[[i, 3]]; 
       dv = blox[[i, 4]]; 
      (10) If[blockoutside2var[uc - du, uc + du, 
                      vc - dv, vc + dv, iter], 
         flag = "drop", 
         If[f[uc, vc] < 0 && inside2var[uc, vc], 
            flag = "fail", 
            If[f[uc, vc] < 0, 
               flag = "subdivide", 
               If[2*f[uc, vc] - Apply[Plus, Abs[Flatten[CoefficientList 
                [f[u*du + uc, v*dv + vc], {u, v}]]]] > 0, 
                  flag = "drop", 
                  flag = "subdivide"]]]] 
      (21) If[flag == "fail", Print["f is negative at (", uc, ",", vc, ")."]]; 
      (22) i = 2; blox = List[]; 
      If[flag == "drop", 
      block = Join[Take[blox, i - 1], Drop[blox, i]]; i = i - 1]; 
      If[flag == "subdivide", 
      hdu = du/2; hdv = dv/2; 
      blox = Join[Take[blox, i - 1], 
      List[List[uc - hdu, vc - hdv, hdu, hdv], 
      List[uc - hdu, vc + hdv, hdu, hdv], 
      List[uc + hdu, vc - hdv, hdu, hdv], 
      List[uc + hdu, vc + hdv, hdu, hdv]], 
      Drop[blox, i]]; i = i + 3]; 
    (33) If[Length[blox] > 0, 
      ListPlot[Map[g, blox], PlotRange -> 
      {{umin, umax}, {vmin, vmax}}]]; 
      Print["Number of blocks remaining: ", Length[blox]]; 
      If[Length[blox] > 0, 
      Print["Width: ", N[2*hdu], " Height: ", N[2*hdv]]] 
    (39) ] ) 

Comments specific to Test2Var:
1. The function g is for the optional plotting of the centers of the remaining blocks (see Line 35).
2. We don’t want to check positivity of $P$ on a whole rectangle, but on a specified domain $R$ inside an initial bounding rectangle. The function blockoutside2var checks whether the current block (sub-rectangle) is completely outside $R$. If so, the flag is set to drop the block.

3. We only set the flag to “fail” if the block center is actually inside $R$.

4. If $P$ is negative and the current block is neither completely outside nor completely inside $R$, we subdivide.

5. The ListPlot command on Line 34 is optional; it plots all the centers of the blocks not yet dropped, giving a pictorial report of the algorithm’s progress.

**Side note:** This algorithm will not terminate if the function $f$ is exactly equal to zero (but never negative) on part of the domain. In such a case we might analyze derivatives, plugged into the algorithm, in the hope of proving that $f \geq 0$ on the domain.

### 7. Inequality tests.

The purpose of this section is to verify the inequalities (7) and (8) given in the proof of Theorem 4.7. Expressions which convert the associated functions to rational polynomials are defined first. Next, the one variable inequality (7) is proven. Then the shape of the domain of the two variable inequality (8) is described so that it can be determined whether or not a given block meets the domain. Finally, the inequality (8) is verified.

#### 7.1. Defined expressions.

We first set up the functions to be tested in the proof of Theorem 4.7. In order to get a polynomial from the trigonometric functions, we start by making the substitutions

\[
\cos \theta = \frac{2u}{1 + u^2}, \quad \sin \theta = \frac{1 - u^2}{1 + u^2}
\]

and

\[
\cos \left( \frac{\alpha}{2} \right) = \frac{2v}{1 + v^2}, \quad \sin \left( \frac{\alpha}{2} \right) = \frac{1 - v^2}{1 + v^2}.
\]

Since these are rational functions, rather than polynomials, the functions $NN$, $nn$ and $RR$ defined below are multiplied by powers of $1 + u^2$ and of $1 + v^2$, in a way that does not affect the final inequalities to be proved.

As $u$ ranges from $-\infty$ up to $\infty$, $\theta$ ranges from $\frac{3\pi}{2}$ down to $-\frac{\pi}{2}$. We will use $u \in [-\frac{3}{2}, \frac{3}{2}]$, which corresponds approximately to $\theta \in [-\frac{\pi}{8}, \frac{9\pi}{8}]$. To get $\alpha \in [0, \frac{\pi}{2}]$ we need $0 \leq \frac{\alpha}{2} \leq \frac{\pi}{4}$, which translates to $1 \geq v \geq \sqrt{2} - 1$. The value $\sqrt{2} - 1$ will be approximated by the slightly smaller value $70/169$.

In order to apply the programs to the inequalities in the proof of Theorem 4.7, we define

\[
\text{CosTheta}[u_] := 2u/(1 + u^2)
\]
\[ \sin \theta_1 := \frac{(1 - u^2)/(1 + u^2)} \]
\[ \cos \alpha := 2v/(1+v^2) \]
\[ \sin \alpha := \frac{(1 - v^2)/(1 + v^2)}{2} \]
\[ \cos \alpha := 2\cos \alpha^2 - 1 \]
\[ \sin \alpha := 2\sin \alpha^2 \cos \alpha^2 \]
\[ \eta[n, u, v] := \text{List} \left\{ \text{Simplify} \left[ \text{Expand} \left[ \left( \cos \alpha \left( 1 - r^2 \right) \cos \theta_1 + \sin \alpha \left( 1 + r^2 \right) \sin \theta_1 \right) \left( 1 - r^2 \right)^2 + \left( 2r \cos \theta_1 \right)^2 \right] \left( 1 + u^2 \right)^3 \left( 1 + v^2 \right)^2 \right] \right\}, \text{Simplify} \left[ \text{Expand} \left[ \left( \cos \alpha \left( 1 - r^2 \right) \sin \theta_1 - \sin \alpha \left( 1 + r^2 \right) \cos \theta_1 \right) \left( 1 - r^2 \right)^2 + \left( 2r \sin \theta_1 \right)^2 \right] \left( 1 + u^2 \right)^3 \left( 1 + v^2 \right)^2 \right] \right\}, 0 \]
\[ \eta[n, u, v] := \text{List} \left\{ -2r \cos \theta_1 \left( 1 + v^2 \right) \left( 1 + u^2 \right), 2r \sin \theta_1 \left( 1 + v^2 \right) \left( 1 + u^2 \right), 0 \right\} \]
\[ \rho[r, u, v] := \text{List} \left\{ \left( 1 + r^2 \right) \sin \alpha \left( 1 + v^2 \right) \left( 1 + u^2 \right), \left( 1 + r^2 \right) \cos \alpha \left( 1 + v^2 \right) \left( 1 + u^2 \right), 0 \right\} \]
\[ \text{fplus}[r, u, v] := \text{Evaluate} \left( \text{Simplify} \left[ \eta[n, u, v] \cdot \rho[r, u, v] \right] \right) \]
\[ \text{fminus}[r, u, v] := \text{Evaluate} \left( \text{Simplify} \left[ \eta[n, u, v] \cdot \rho[r, u, v] \right] \right) \]
\[ G[v] := -19/64 + 7/3 \cos \alpha \left( 1 - r^2 \right) \]
\[ \text{outsideleft2var}[u, v] := \text{If} \left[ G[v] < \cos \theta_1, \text{True}, \text{False} \right] \]
\[ \text{outside2var}[u, v] := \text{If} \left[ G[v] > \sin \theta_1 \right] \]
\[ \text{r0} := 7704/10000 \]
\[ \text{f1}[u, v] := \text{fminus}[r0, u, v] \]
\[ \text{f2}[u, v] := \text{fplus}[r0, u, v] \]

Note that the “Evaluate” command in the definition of fplus and fminus causes the “Simplify” command to occur immediately instead of every time the functions are called.

7.2. Verifying the one variable inequality: Establishing the range for the 2-variable inequality. The following lemma verifies the one variable inequality (7) in the proof of Theorem 4.7:
Lemma 7.1.

\[ \tan^{-1} \left[ \frac{1 + r^2}{1 - r^2} \tan \alpha \right] \leq \omega(\alpha) \]

and

\[ \omega(\alpha) \leq \tan^{-1} \left[ \frac{1 + r^2}{1 - r^2} \tan \alpha \right] + \frac{\pi}{2} \]

for \( \alpha \in [0, \frac{\pi}{2}] \), with the inverse tangent defined to equal \( \frac{\pi}{2} \) when \( \alpha = \frac{\pi}{2} \).

Proof. The first inequality is a little delicate to prove near \( \alpha = \frac{\pi}{2} \). That is because in order to convert everything to rational functions, we take a cosine and then square both sides. Even though the original inequality is true on the whole interval, there are points near \( \frac{\pi}{2} \) for which the squared inequality is false.

To remedy this, we first note that \( \tan^{-1} \left[ \frac{1 + r^2}{1 - r^2} \tan \alpha \right] \in [0, \frac{\pi}{2}] \) for all \( \alpha \in [0, \frac{\pi}{2}] \).

We will divide \( [0, \frac{\pi}{2}] \) into two intervals \( [0, a_0] \) and \( [a_0, \frac{\pi}{2}] \), such that the squared inequality is true for \( \alpha \in [0, a_0] \), and such that \( \omega(\alpha) > \frac{\pi}{2} \) for \( \alpha \in [a_0, \frac{\pi}{2}] \). Each of these last inequalities implies the original (1).

Let \( a_0 = \cos^{-1} \left( \frac{41}{841} \right) \). With the rationalizing substitution \( \cos \frac{\alpha}{2} = \frac{2v}{1 + v^2} \), \( \alpha = a_0 \) corresponds to \( v = 3/7 \). We need to check that \( -\cos(\omega(\alpha)) > 0 \) for \( \alpha \in [a_0, \frac{\pi}{2}] \), that is, for \( v \in [\sqrt{2} - 1, 3/7] \). To get rational endpoints, we will actually prove the inequality for \( v \in [2/5, 3/7] \). This is done by the command

\[ \text{Test1Var[Poly2G, 2/5, 3/7]} \]

where

\[ \text{Poly2G}[v] = -G[v] * (1 + v^2)^2 \]

\[ = -\frac{1}{960}(-957 + 4480v - 7610v^2 + 4480v^3 + 2307v^4) \]

The test comes back successful.

We now need to prove the “squared inequality” for \( \alpha \in [0, a_0] \):

\[ \tan^{-1} \left[ \frac{1 + r^2}{1 - r^2} \tan \alpha \right] \leq \omega(\alpha) \]

\[ \cos \left( \tan^{-1} \left[ \frac{1 + r^2}{1 - r^2} \tan \alpha \right] \right) \geq \cos(\omega(\alpha)) = G[v] \]

\[ \cos^2 \left( \tan^{-1} \left[ \frac{1 + r^2}{1 - r^2} \tan \alpha \right] \right) \geq G[v]^2 \]
\[
\frac{1}{\left[ \frac{1+r^2}{1-r^2} \tan \alpha \right]^2 + 1} \geq G[v]^2
\]

\[
1 \geq \left( \frac{1+r^2}{1-r^2} \tan \alpha \right)^2 + 1 \quad G[v]^2
\]

\[
1 - \left( \left( \frac{1+r^2}{1-r^2} \tan \alpha \right)^2 + 1 \right) G[v]^2 \geq 0
\]

\[
1 - \left( \frac{1+r^2}{1-r^2} \tan \alpha \right)^2 + 1 \right) G[v]^2 \geq 0.
\]

Setting \( r = \frac{7704}{10000} \) and multiplying both sides by \((v^2 + 1)^4(v^4 - 6v^2 + 1)^2\) turns the left side into a polynomial which we name MainPoly1[v] and invoke the command

\[
\text{Test1Var[MainPoly1,3/7,1]}
\]

which comes back successful.

To prove inequality (2) above, we first show that \( G[v] \) is positive for \( v \in [1/2, 1] \), so that \( \cos^{-1}(\omega(\alpha)) < \frac{\pi}{2} \), which implies (2) for the corresponding values of \( \alpha \). This is done by defining

\[
\text{Poly1G} = G[v]* (1+v^2)^2 = \frac{1}{960}(-957+4480v-7610v^2+4480v^3+2307v^4).
\]

We invoke the command

\[
\text{Test1Var[Poly1G,1/2,1]}
\]

which comes back successful.

We then show that for \( v \in [2/5, 1/2] \),

\[
G[v] \geq \cos \left( \tan^{-1} \left( \frac{1+r^2}{1-r^2} \frac{4v(v^2 - 1)}{v^4 - 6v^2 + 1} \right) + \frac{\pi}{2} \right).
\]

Here the right side is negative, so the above inequality would follow from

\[
G[v]^2 \leq \cos^2 \left( \tan^{-1} \left( \frac{1+r^2}{1-r^2} \frac{4v(v^2 - 1)}{v^4 - 6v^2 + 1} \right) + \frac{\pi}{2} \right)
\]

\[
= \sin^2 \left( \tan^{-1} \left( \frac{(1+r^2)(4v)(v^2 - 1)}{(1-r^2)(v^4 - 6v^2 + 1)} \right) \right)
\]

\[
= \left[ \frac{(1+r^2)(4v)(v^2 - 1)}{(1-r^2)(v^4 - 6v^2 + 1)} \right]^2 + \frac{1}{1}
\]

\[
\text{Test1Var[Poly1G,1/2,1]}
\]

which comes back successful.
Hence
\[
\left( \left( \frac{(1 + r^2)(4v)(v^2 - 1)}{(1 - r^2)(v^4 - 6v^2 + 1)} \right)^2 + 1 \right) G[v]^2 \leq \left( \frac{(1 + r^2)(4v)(v^2 - 1)}{(1 - r^2)(v^4 - 6v^2 + 1)} \right)^2,
\]
\[
\left( (1 + r^2)(4v)(v^2 - 1) \right)^2 + \left( (1 - r^2)(v^4 - 6v^2 + 1) \right)^2 \geq (1 + r^2)(4v)(v^2 - 1),
\]
and
\[
\left( (1 + r^2)(4v)(v^2 - 1) \right)^2 - \left( (1 + r^2)(4v)(v^2 - 1) \right)^2 + \left( (1 - r^2)(v^4 - 6v^2 + 1) \right)^2 \geq 0.
\]

Multiplying by \((1 + v^2)^4\) to make the above a polynomial, we test it for \(v \in [2/5, 1/2]\). Test1Var immediately confirms to be positive. 

### 7.3. Establishing the shape of the domain.

Now to justify the test called “blockoutside2var” in the program “Test2Var,” we need to show that if the two right-hand corners of a rectangle are to the left of \(R\), or if the two left-hand corners of the rectangle are to the right of \(R\), then the whole rectangle is outside \(R\). The top and bottom of \(R\) are the straight lines \(v = \frac{70}{169}\) and \(v = 1\), and the top and bottom of a subrectangle will always be in \([\frac{70}{169}, 1]\). We will show that the left and right sides of \(R\) are formed by functions \(u_1(v) < u_2(v)\) which are both increasing functions, from which the desired result will follow. We start with the following lemma:

**Lemma 7.2.** The function \(G[v]\) is increasing on \([\frac{70}{169}, 1]\).

**Proof.** We take the first derivative \(G'(v)\) and multiply it by \((1 + v^2)^3\), obtaining the polynomial
\[
\frac{14}{3} - \frac{178}{15}v + \frac{382}{15}v^3 - \frac{14}{3}v^4.
\]
To show this is positive on \([\frac{70}{169}, 1]\), we call it \(GP[v]\) and invoke the command
\[
\text{Test1Var}[\text{GP}, \frac{70}{169}, 1]
\]
which is successful. 

**Lemma 7.3.** Within the rectangle \((u, v) \in \left[\frac{33}{22}\right] \times \left[\frac{70}{169}, 1\right]\), the expressions
\[
(1) \quad \frac{1 - u^2}{1 + u^2} \geq G(v) \geq \frac{2u}{1 + u^2}
\]
and
\[
(2) \quad \frac{2u}{1 + u^2} = G(v)
\]
implicitly define single-valued, increasing functions \( u = u_1(v) \) and \( u = u_2(v) \), respectively, with \( u_1(v) < u_2(v) \).

**Proof.** First, calculations show that \( G[\frac{70}{169}] > -0.016 \) and \( G[1] < 0.704 \). Since \( G \) is increasing, its values on all of \( [\frac{70}{169}, 1] \) are between these endpoint values.

Now it is straightforward to show that \( 2u/(1+u^2) > 0.9 \) on \( [\frac{3}{4}, \frac{3}{2}] \) and \( 2u/(1+u^2) < -0.4 \) on \( [-\frac{3}{2}, -\frac{1}{4}] \). So, given the range of \( G[v] \), Equation (2) will only allow values of \( u \) within \( [-\frac{1}{4}, \frac{3}{4}] \). On this interval, the function \( 2u/(1+u^2) \) is strictly increasing. Taking these facts together, we see that there will be exactly one \( u \) for each \( v \in [\frac{70}{169}, 1] \) for which (2) holds, and that the function \( u = u_2(v) \) thereby defined is a strictly increasing function.

Next, within \( [-\frac{3}{2}, \frac{3}{2}] \) the inequality \( (1-u^2)/(1+u^2) \geq 2u/(1+u^2) \) is true precisely for \( u \in [-\sqrt{2}, 1-\sqrt{2}] \). But for \( u \in (\sqrt{2}-1, \sqrt{2}-1], (1-u^2)/(1+u^2) \geq \sqrt{2}/2 \), and so \( (1-u^2)/(1+u^2) \) is outside the range of \( G[v] \). Thus, for (1) to hold, \( u \) must be within \( [-\frac{3}{2}, 1-\sqrt{2}] \). On this interval, \( (1-u^2)/(1+u^2) \) is strictly increasing, and again we obtain a single-valued, increasing function \( u = u_1(v) \) implicitly defined by (1).

Finally, since \( -1/4 \leq u_2 \leq 3/4 \) and \( -3/2 \leq u_1 \leq 1 - \sqrt{2} < -1/4 \), we see that \( u_1(v) < u_2(v) \).

From the above lemma, we infer that the shape of \( R \) is such that the following proposition holds:

**Proposition 7.4.** If the \( v \)-values at the top and bottom of a rectangle \( B \) are within \( [\frac{70}{169}, 1] \), and if either the two right corners of \( B \) are to the right of \( R \) or the two left corners of \( B \) are to the right of \( R \), then the entire rectangle is outside \( R \).

Finally, we use the following lemma to justify the tests “outsideright2var” and “outsideleft2var”. In the lemma \( K \) represents \( G[v] \) which does lie between \( -\sqrt{2}/2 \) and \( \sqrt{2}/2 \), as noted previously.

**Lemma 7.5.** If

\[
-\frac{\sqrt{2}}{2} \leq K \leq \frac{\sqrt{2}}{2}
\]

and

\[
\cos^{-1} K \leq \theta \leq \cos^{-1} K + \frac{\pi}{2}
\]

then

\[
\cos \theta \leq K \leq \sin \theta.
\]

**Proof.** First note that \( \frac{\pi}{4} \leq \cos^{-1} K \leq \frac{3\pi}{4} \), and thus \( \frac{\pi}{4} \leq \theta \leq \frac{5\pi}{4} \). In particular, \( \cos^{-1} K \in [0, \pi] \), where the cosine function is decreasing. To prove that \( \cos \theta \leq K \), consider
Case a1: $\theta \in [0, \pi]$. The fact that $\cos^{-1} K \leq \theta$ implies $K = \cos(\cos^{-1} K) \geq \cos \theta$.

Case a2: $\theta \in [\pi, \frac{5\pi}{4}]$. Then $\cos \theta \leq -\frac{\sqrt{2}}{2} \leq K$.

Now to prove that $K \leq \sin \theta$, consider:

Case b1: $\theta \in [\frac{\pi}{4}, \frac{3\pi}{4}]$. Then $K \leq \sqrt{\frac{2}{2}} \leq \sin \theta$.

Case b2: $\theta \in [\frac{3\pi}{4}, \frac{5\pi}{4}]$. Then both $\theta$ and $\cos^{-1} K + \frac{\pi}{2}$ are in an interval where the sine function is decreasing. So the fact that $\theta \leq \cos^{-1} K + \frac{\pi}{2}$ implies that $\sin \theta \geq \sin(\cos^{-1} K + \frac{\pi}{2}) = K$, as desired. \hfill $\Box$

7.4. The 2-variable inequality. The polynomials $f_1(u, v)$ and $f_2(u, v)$ are the functions we need to prove positive on $\sqrt{2} - 1 \leq v \leq 1$, $G(v) \leq u \leq G(v) + \frac{\pi}{2}$ in order to verify the inequality (8) in the proof of Theorem 4.7. The following two commands accomplish this:

$$\text{Test2Vars}[f_1, -\frac{3}{2}, \frac{3}{2}, \frac{70}{169}, 1]$$
$$\text{Test2Vars}[f_2, -\frac{3}{2}, \frac{3}{2}, \frac{70}{169}, 1].$$

The test comes back positive.

References


Received April 4, 2002.

DEPARTMENT OF MATHEMATICS
BRIGHAM YOUNG UNIVERSITY
PROVO, UTAH 84602
E-mail address: mdorff@math.byu.edu

DEPARTMENT OF MATHEMATICS
BRIGHAM YOUNG UNIVERSITY
PROVO, UTAH 84602
E-mail address: deniseh@math.byu.edu

DEPARTMENT OF MATHEMATICS
BRIGHAM YOUNG UNIVERSITY
PROVO, UTAH 84602
E-mail address: lawlor@mathed.byu.edu
HILBERT’S TENTH PROBLEM FOR ALGEBRAIC
FUNCTION FIELDS OF CHARACTERISTIC 2

Kirsten Eisenträger

Let $K$ be an algebraic function field of characteristic 2 with constant field $C_K$. Let $C$ be the algebraic closure of a finite field in $K$. Assume that $C$ has an extension of degree 2. Assume that there are elements $u, x$ of $K$ with $u$ transcendental over $C_K$ and $x$ algebraic over $C(u)$ and such that $K = C_K(u, x)$. Then Hilbert’s Tenth Problem over $K$ is undecidable. Together with Shlapentokh’s result for odd characteristic this implies that Hilbert’s Tenth Problem for any such field $K$ of finite characteristic is undecidable. In particular, Hilbert’s Tenth Problem for any algebraic function field with finite constant field is undecidable.

1. Introduction.

Hilbert’s Tenth Problem in its original form can be stated in the following form: Is there a uniform algorithm that determines, given a polynomial equation with integer coefficients, whether the equation has an integer solution or not? In [Mat70] Matijasevich proved that the answer to this question is no, i.e., that Hilbert’s Tenth Problem is undecidable. Since then various analogues of this problem have been studied by asking the same question as above for polynomial equations with coefficients and solutions over some other commutative ring $R$. Perhaps the most important unsolved question in this area is Hilbert’s Tenth Problem over the field of rational numbers. There are also many results that prove undecidability: It was proved in [Den80] and [DL78] that Hilbert’s Tenth Problem is undecidable for various rings of algebraic integers, and [Den78] proves the undecidability of the problem for rational functions over formally real fields. In [KR92a] Kim and Roush proved that Hilbert’s Tenth Problem over $\mathbb{C}(t_1, t_2)$ is undecidable. Diophantine undecidability has also been proved for some rational function fields of characteristic $p$: Pheidas [Phe91] has shown that Hilbert’s Tenth Problem is undecidable for rational function fields over finite fields of characteristic greater than 2 and Videla [Vid94] has proved the analogous result for characteristic 2. Kim and Roush [KR92b] proved undecidability for rational function fields of characteristic $p > 2$ whose constant fields do not contain the algebraic closure of a finite field. In [Shl00] Shlapentokh
proved that the problem for algebraic function fields over possibly infinite constant fields of characteristic $p > 2$ is undecidable. This paper will solve the analogous problem over function fields of characteristic 2, so Hilbert’s Tenth Problem for any such field of finite characteristic is undecidable. We will first describe the general approach that is used to prove the undecidability of Hilbert’s Tenth Problem for any function field of positive characteristic. The approach is based on an idea that was first introduced by Denef in [Den79] and further developed by Pheidas in [Phe91] and Shlapentokh in [Shl96] and [Shl00].

Before we can describe the idea in detail we need to define what an algebraic function field is:

**Definition 1.1.** A field extension $K/C_K$ is said to be an algebraic function field (of one variable) if these conditions hold:

1) The transcendence degree of $K/C_K$ is 1;
2) $K$ is finitely generated over $C_K$; and
3) $C_K$ is algebraically closed in $K$.

In this case there exists $t \in K$, transcendental over $C_K$, such that the degree of the field extension $[K : C_K(t)]$ is finite. The field $C_K$ is called the constant field of $K$.

We also need to define two notions that we will use below:

**Definition 1.2.**

1. If $R$ is a commutative ring, a diophantine equation over $R$ is an equation $P(x_1, \ldots, x_n) = 0$ where $P$ is a polynomial in the variables $x_1, \ldots, x_n$ with coefficients in $R$.
2. A subset $S$ of $R^k$ is diophantine if there is a polynomial $P(x_1, \ldots, x_k, y_1, \ldots, y_m) \in R[x_1, \ldots, x_k, y_1, \ldots, y_m]$ such that

$$S = \{(x_1, \ldots, x_k) \in R^k : \exists y_1, \ldots, y_m \in R, (P(x_1, \ldots, x_k, y_1, \ldots, y_m) = 0)\}.$$  

When $R$ is not a finitely generated algebra over $\mathbb{Z}$, we restrict our attention to diophantine equations whose coefficients are in a finitely generated algebra over $\mathbb{Z}$. In particular, if $R$ is a ring of polynomials or a field of rational functions in an indeterminate $t$, we only consider diophantine equations whose coefficients lie in the natural image of $\mathbb{Z}[t]$ in $R$.

**1.1. Idea of Proof.** Let $\mathbb{N}$ be the set of natural numbers $\{0, 1, 2, \ldots \}$. The general idea of the proof is to reduce a certain decision problem over the natural numbers which we know to be undecidable to Hilbert’s Tenth Problem over $K$. The undecidable structure that we will use is the diophantine theory of the natural numbers with addition and a predicate $|_p$ defined by $n|_p m$ if and only if $\exists \in \mathbb{N}(m = p^n)$. In [Phe87] Pheidas showed that this structure has an undecidable diophantine theory, i.e., there is no uniform algorithm that, given a system of equations over the natural numbers with
addition and $|_p$, determines whether this system has a solution or not. To reduce this problem to Hilbert’s Tenth Problem over $K$ we first let $G$ be a subfield of $K$ containing an element $t$ transcendental over $C_K$. The field $G$ will be defined in Lemma 2.2. Also fix a prime $p$ of $K$ which lies above a non-trivial prime of $G$. We can choose $t$ and $p$ such that $\text{ord}_p t = 1$. Both $t$ and $p$ will be defined at the end of Section 2. Let $O_{K,p} := \{ x \in K : \text{ord}_p x \geq 0 \}$, and let $O_{G,p} := G \cap O_{K,p}$. Now let $\text{INT}(p)$ be any subset of $K$ such that $O_{G,p} \subseteq \text{INT}(p) \subseteq O_{K,p}$. We define a map $f$ from the integers to subsets of $K$ by associating to an integer $n$ the subset $f(n) := \{ x \in \text{INT}(p) : \text{ord}_p x = n \}$. Then $n_3 = n_1 + n_2$ ($n_i \in \mathbb{N}$) is equivalent to the existence of $z_i \in f(n_i)$ such that $z_3 = z_1 \cdot z_2$. This follows from the fact that $\text{ord}_p z_1 + \text{ord}_p z_2 = \text{ord}_p (z_1 \cdot z_2)$ and that $t^{n_1} \in f(n_i)$. We also have for natural numbers $n, m$

$$n|_p m \iff \exists s \in \mathbb{N} \ m = p^s n$$
$$\iff \exists x \in f(n) \ \exists y \in f(m) \ \exists s \in \mathbb{N} \ (\text{ord}_p y = p^s \text{ord}_p x).$$

This equivalence can be seen easily, because we can let $x := t^n$ and $y := t^m$. But the last formula is equivalent to

$$\exists x \in f(n) \ \exists y \in f(m) \ \exists w \in K \ \exists s \in \mathbb{N} \ w = x^{p^s} \text{ and } \{ w/y, y/w \} \subset \text{INT}(p).$$

Here $w/y \in \text{INT}(p)$ and $y/w \in \text{INT}(p)$ just means that $y$ and $w$ have the same order at $p$.

If we have diophantine definitions for $p(K) := \{(x, w) \in K^2 : \exists s \in \mathbb{N}, w = x^{p^s}\}$ and $\text{INT}(p)$, then the above argument shows that for every system of equations with addition and $|_p$ we can construct a system of polynomial equations over $K$ which will have solutions in $K$ if and only if the original system of equations over $\mathbb{N}$ has solutions in $\mathbb{N}$. But the diophantine theory of $\mathbb{N}$ with $+$ and $|_p$ is undecidable; hence Hilbert’s Tenth Problem over $K$ is undecidable.

So the strategy for the proof will be to prove that $p(K)$ is diophantine and that there exists some set $\text{INT}(p)$ as above which is diophantine for the class of fields $K$ that we are considering. This can be summarized as:

**Theorem 1.3.** Let $K$ be an algebraic function field of characteristic 2 with constant field $C_K$. Let $C$ be the algebraic closure of a finite field in $K$. Assume that $C$ has an extension of degree 2. Assume that there are elements $u, x$ of $K$ with $u$ transcendental over $C_K$ and $x$ algebraic over $C(u)$ and such that $K = C_K(u, x)$. Then $p(K)$ is diophantine. Also there exists a subfield $G$ of $K$ as above with $C(t) \subseteq G$ for an element $t$ transcendental over $C_K$. There exists a prime $p$ of $K$ satisfying the conditions above such that $\text{INT}(p)$ is diophantine for some set $\text{INT}(p)$ with $O_{G,p} \subseteq \text{INT}(p) \subseteq O_{K,p}$. So Hilbert’s Tenth Problem over $K$ is undecidable.

In [Shl00] Shlapentokh proves that for such $K$ in any characteristic $p > 0$ there exists some set $\text{INT}(p)$ as above which is diophantine. She also proves
that \( p(K) \) is diophantine when the characteristic of \( K \) is greater than 2, but her main lemmas are not valid in characteristic 2. So in order to prove undecidability in characteristic 2, the last open case, we need to prove that \( p(K) \) is diophantine when the characteristic of \( K \) is 2. The rest of the paper is devoted to proving this. The outline of the proof follows Shlapentokh’s proof for odd characteristic. Before we can prove this we first need to prove some properties of \( K \) and then set up some notation. The next section will do that. In Section 3 we will prove that the set \( p(K) \) is diophantine in characteristic 2.

2. Setup and notation.

Let \( \mathbb{N} \) be the set of natural numbers \( \{0, 1, 2, \ldots \} \). Let \( K, C_K, C, u \) and \( x \) be as in Theorem 1.3. We will use the following:

**Notation 2.1.** Let \( F \) be a field, and \( k \in \mathbb{N} \). We denote by \( F^k := \{a^k : a \in F \} \).

We will now prove some properties of \( K \) that we will need later on. We may assume that \( u \) is not a square in \( K \), because if \( u = u_1^2 \) with \( u_1 \in K \) and \( s \in \mathbb{N} \), we can replace \( u \) by \( u_1 \). Then \( K = C_K(u, x) = C_K(u_1, x) \). Since the extension \( K/C_K(u) \) can be generated by a single element, \( u \in K^2 \) only if \( s \leq [K : C_K(u)] \), so replacing \( u \) by its square root terminates after a finite number of steps.

We have the following:

**Lemma 2.2.** Let \( K, C_K, C, u, x \) be as above. Let \( G \) be the algebraic closure of \( C(u) \) inside \( K \). Then \( G = C(u, x) \).

**Proof.** First note that \( C(u) \) is algebraically closed in \( C_K(u) \), because \( C \) is algebraically closed in \( C_K (\cite{Deu73}, p. 117) \). Let \( m := [K : C_K(u)] \). If \( m = 1 \), i.e., \( x \in C_K(u) \), then the statement is true since \( C(u) \) is algebraically closed in \( C_K(u) \). So assume \( x \notin C_K(u) \). Let \( \alpha \in G, \alpha \notin C(u) \). Then by \( \cite{Lan93} \) Lemma 4.10, p. 366,

\[
(1) \quad [C(u, \alpha) : C(u)] = [C_K(u, \alpha) : C_K(u)] \leq [K : C_K(u)] = m.
\]

In particular, \( [C(u, x) : C(u)] = m \). Now assume by contradiction that there exists a \( \beta \in G, \beta \notin C(u, x) \). Let \( G_1 := C(u, x, \beta) \). Then \( [G_1 : C(u)] > m \).

Also \( G_1 \) is an algebraic function field with constant field \( C \), and \( C \) is perfect. Then by \( \cite{Mas84} \), p. 94 the extension \( G_1/C(u) \) is finite and separable, since \( u \) is not a square in \( G_1 \). Hence there exists a primitive element \( \gamma \in G_1 \) with \( C(u, \gamma) = G_1 \). But then \( [C(u, \gamma) : C(u)] = [G_1 : C(u)] > m \), contradicting (1).

**Definition 2.3.** Let \( K \) be an algebraic function field with constant field \( C_K \). A **constant field extension of \( K \)** is an algebraic function field \( L \) with
constant field $C_L$ such that $L \supseteq K$, $C_L \cap K = C_K$ and $L$ is the composite extension of $K$ and $C_L$, $L = C_L K$.

**Proposition 2.4.** Let $G$ be as in Lemma 2.2. Fix a positive integer $k$. For any sufficiently large positive integer $h$ a finite constant extension of $G$ contains a nonconstant element $t$ and a set of constants $V$ of cardinality $k + 2^h$ such that $0 \in V$, $1 \notin V$. Also we can choose $t$ and $V$ such that for all $c \in V$ the divisor of $t + c$ is of the form $p_c/q$, where the $p_c$'s and $q$ are prime divisors of degree $2^h$.

**Proof.** This is Theorem 6.11 of [Shl00] if $C$ is infinite. The proof of the existence of $t$ and $V$ with the desired properties in Theorem 6.11 does not use that $C$ is infinite; it only requires passing to a finite extension of $C$. □

**Remark.** In Proposition 2.4 we can choose $V$ with the property that for all $s \in \mathbb{N}$ for all $c, c' \in V$ $c^s \neq c'$ if $c \neq c'$.

From now on we will assume that an element $t$ and a set $V$ of constants with the desired properties as in Proposition 2.4 already exist in $G$. (Otherwise rename the constant extension $G$ again and work with it instead.) Enlarging the field of constants by a finite extension is okay as far as the undecidability of Hilbert’s Tenth Problem is concerned. Also let $p := p_0$, so that the divisor of $t$ is of the form $p/q$.

**Proposition 2.5.** Let $G, C, t$ be as above. Then $[G : C(t)]$ is separable, and $2^h = n = [G : C(t)]$.

**Proof.** Since the divisor of $t$ is of the form $p/q$, $t$ is not a square in $G$. Also $C$ is perfect. Hence $G/C(t)$ is separable by [Mas84], p. 22. Also by [FJ86], p. 13, $[G : C(t)] = \deg p = \deg q = 2^h$. □

Now we can prove that $K$ is separably generated:

**Corollary 2.6.** Let $K$ be an algebraic function field with constant field $C_K$. Let $C$ be the algebraic closure of a finite field in $K$. Assume that $C$ has an extension of degree $2$. Assume that there exist $x, u$ as above. Let $G, t$ be as above. Then $K/C_K(t)$ is separable.

**Proof.** By Proposition 2.5 $G/C(t)$ is separable. The field $K$ is the compositum of $C_K(t)$ and $G$ over $C(t)$, hence $K/C_K(t)$ is also separable. □

Now we can use Lemma 6.13 of [Shl00] to see how the $p_c$'s and $q$ behave in the extension $K$.

**Lemma 2.7** (Lemma 6.13 of [Shl00]). Let $H$ be an algebraic function field over a field of constants $C_H$. Let $K$ be a constant field extension of $H$. Let $C_K$ be the constant field of $K$, and assume $H$ is algebraically closed in $K$. Let $t \in H - C_H$ be such that $H/C_H(t)$ is separable. Let $a$ be a prime of $C_H(t)$ remaining prime in the extension $H$ and such that its residue field is separable over $C_H$. Then $a$ will have just one prime factor in $K$. 

This lemma easily implies the following corollary:

**Corollary 2.8.** Let \( \{p_c : c \in V\} \) and \( q \) be as in Proposition 2.4. Then the \( p_c \)'s and \( q \) remain prime in \( K \).

**Proof.** Lemma 2.7 applies, since \( K/C_K \) is a constant field extension of \( G/C \): By construction \( C \) is algebraically closed in \( C_K \), and also \( C_K G = K \). The only thing we need to check is that \( G \cap C_K = C \). Assume \( \alpha \in G \setminus C \). Then \( \alpha \) is transcendental over \( C \) and also over \( C_K \). Hence \( \alpha \notin G \cap C_K \). Thus we can apply the lemma to the primes, \( t, 1/t \) and \( t + c \) of \( C(t) \). Since \( C \) is perfect, the residue extensions of the primes will be separable. \( \square \)

Since the \( p_c \)'s and \( q \) remain prime in \( K \) we will just denote them by the same letters again when considering them as primes of \( K \), and we will let \( p := p_0 \). Now we can fix some notation that we will use for the rest of the paper:

- \( K \) will denote an algebraic function field over a field of constants \( C_K \) of characteristic \( p = 2 \).
- \( C \) will denote the algebraic closure of a finite field inside \( C_K \).
- \( t \) will denote a nonconstant element of \( K \setminus C_K \) such that the divisor of \( t \) is of the form \( p/q \), where \( p, q \) are primes of degree \( 2^h \) for some natural number \( h \). Furthermore, \( K/C_K(t) \) is separable, and \( 2^h = n = [K : C_K(t)] \).
- \( \overline{C}_K \) will denote the algebraic closure of \( C_K \), and \( \overline{K} := C_K \).
- \( r \) will denote the number of primes of \( \overline{K} \) ramifying in the extension \( \overline{K}/C_K(t) \).
- \( V \) will denote a subset of \( C \), containing \( n + 2r + 6 \) elements, such that \( 0 \in V, 1 \notin V \), and such for all \( c \in V \) the divisor of \( t + c \) is of the form \( p/cq \), where \( p_c \) is a prime divisor of \( K \). Also pick \( V \) such that for any \( s \in \mathbb{N}, c, c' \in V \), we have \( c^s \neq c' \), if \( c \neq c' \).
- For all \( c \in V, \mathfrak{P}_c \) will denote the prime of \( C_K(t) \) lying below \( p_c \), while \( \mathfrak{Q} \) will denote the prime of \( C_K(t) \) lying below \( q \). Also let \( \mathfrak{P} := \mathfrak{P}_0 \). For all \( c \in V, \mathfrak{P}_c \) and \( \mathfrak{Q} \) do not split in the extension \( K/C_K(t) \).
- For every \( c \in V, V_c \) will denote the set \( V_c := \{ c^j : j \in \mathbb{N} \} \). Since every \( c \in V \) is algebraic over a finite field, \( V_c \) is a finite set for all \( c \in V \).

To obtain \( t \) and \( V \) with the desired properties, we have to assume that \( C \) is sufficiently large, but this is not a restriction because we can enlarge the field of constants and by Proposition 2.4 a finite extension is enough. Let \( L \) be this finite extension. If Hilbert’s Tenth Problem over \( L \) is undecidable, then Hilbert’s Tenth Problem over \( K \) is also undecidable. So in the following we will assume that \( L = K \) to simplify notation.
3. \(p\)-th power equations.

Using the notation that we set up in the last section will now prove that the set \(p(K) = \{(x,y) \in K^2 : \exists s \in \mathbb{N}, y = x^s\}\) is diophantine which is Theorem 3.12 below. The main ingredient for proving this is the next theorem. It gives an equivalent definition of what it means for \((x,y)\) to be in \(p(K)\). Eventually we want to find polynomial equations describing these relations, so the goal afterwards will be to rewrite the equations below as polynomial equations.

**Theorem 3.1.** Given \(x, y \in K\), let \(u := \frac{x^2 + t^2 + t}{x^2 + t}\) and \(\bar{v} := \frac{x^2 + t^{-2} + t^{-1}}{x^2 + t^2}\). Let \(v := \frac{y^2 + t^{2i+1} + t^2}{y^2 + t^2}\) and \(\bar{v} := \frac{y^2 + t^{-2i+1} + t^{-2}}{y^2 + t^{-2}}\) for some \(s \in \mathbb{N}\).

Then \(y = x^s\) if and only if

\[
\begin{align*}
(2) & \quad \exists r \in \mathbb{N} \; v = u^r \\
(3) & \quad \exists j \in \mathbb{N} \; \bar{v} = \bar{u}^{2j}.
\end{align*}
\]

**Proof.** Suppose \(y = x^s\). Let \(r = j = s\). Then (2) and (3) are satisfied. This completes one direction of the proof.

On the other hand, suppose that \(r\) and \(j\) as in the statement of the theorem exist. Then

\[
v = \left(\frac{x^2 + t^2 + t}{x^2 + t}\right)^{2r} = \frac{x^{2r+1} + t^{2r+1} + t^{2r}}{x^{2r+1} + t^{2r}} = \frac{y^2 + t^{2s+1} + t^{2s}}{y^2 + t^{2s}}.
\]

So

\[
(x^{2r+1} + t^{2r+1} + t^{2r})(y^2 + t^{2s}) = (x^{2r+1} + t^{2s})(y^2 + t^{2s+1} + t^{2s}),
\]

i.e.,

\[
t^{2r+1}y^2 + t^{2r+1+2s} = x^{2r+1}t^{2s+1} + t^{2r+2s+1}.
\]

Thus

\[
y^2 = (x^{2r+1}t^{2s+1} + t^{2r+2s+1} + t^{2r+1+2s}) \cdot t^{-2r+1}.
\]

Hence if we can show that \(r = s\), then \(y^2 = x^{2s+1}\), so \(y = x^s\), since the characteristic of \(K\) is 2. So our goal is to show that \(r = s\).

Similarly to the calculations above we get

\[
\bar{v} = \left(\frac{x^2 + t^{-2} + t^{-1}}{x^2 + t^{-1}}\right)^{2j} = \frac{x^{2j+1} + t^{-2j+1} + t^{-2j}}{x^{2j+1} + t^{-2j}} = \frac{y^2 + t^{-2s+1} + t^{-2s}}{y^2 + t^{-2s}},
\]

and we get

\[
y^2 = (x^{2j+1}t^{-2s+1} + t^{-2j-2s+1} + t^{-2j+1-2s}) \cdot t^{2j+1}.
\]

By (4)

\[
y = (x^{2r}t^2 + t^{2r-1+2s} + t^{2r+2s-1}) \cdot t^{-2r}.
\]
(unless \( r \) or \( s \) are \(< 1 \)), and by (5)

\[
y = (x^{2^j} t^{-2^s} + t^{-2^s - 2^j - 1} + t^{-2^j - 2^{s - 1}}) \cdot t^{2^j}
\]

(unless \( j \) or \( s \) \(< 1 \)). Eliminating \( y \) from (6) and (7), we get

\[
(t^{2^j - 2^r} x^{2^r}) + (t^{2j - 2^s} x^{2^j}) = t^{2^{s - 2r - 1}} + t^{2j - 1 - 2^s} + t^{-2^{s - 1}}.
\]

Now assume that \( y \) is a square, say \( y = z^2 \) (and \( s, j, r > 0 \)). Then

\[
v = \frac{z^4 + t^{2^s + 1} + t^{2^s}}{z^4 + t^{2^r}} = \left( \frac{z^2 + t^{2^s} + t^{2^{s - 1}}}{z^2 + t^{2^{r - 1}}} \right)^2 = (v')^2.
\]

Hence

\[
v = (v')^2 = u^{2^r}, \text{ so } u^{2^{r - 1}} = \left( \frac{z^2 + t^{2^s} + t^{2^{s - 1}}}{z^2 + t^{2^{r - 1}}} \right) = v'.
\]

Similarly \( \tilde{v} = (\tilde{v}')^2 \) and \( \tilde{u}^{2^{j - 1}} = \tilde{v}' \), so in the new formulae \( s, r \) and \( j \) are replaced by \( s - 1, r - 1 \) and \( j - 1 \), respectively, and we’re done if we can show that \( z = x^{2^{s - 1}} \). Hence we can reduce the problem to the case where either (a) \( s = 0 \) or \( r = 0 \) or \( j = 0 \), or (b) \( y \) is not a square.

**Case (a).** \( s = 0 \): If \( s = 0 \), then \( v = \frac{y^2 + t^2 + t}{y^2 + t} \), and \( v \) is not a square since

\[
\frac{dv}{dx} = \frac{y^2 + t^2 + t + y^2 + t}{(y^2 + t)^2} = t^2 \neq 0.
\]

So if \( s = 0 \), then \( v = \frac{y^2 + t^2 + t}{y^2 + t} = u^r \). Since \( v \) is not a square, this implies \( r = 0 \). Hence \( r = s = 0 \) and we’re done.

If \( r = 0 \), then \( v = u \). By the same argument as above \( u \) is not a square. Now if \( s > 0 \), then \( v \) is a square and hence \( u \) is a square, contradiction. Hence \( r = s = 0 \), and we’re done. The case \( j = 0 \) follows from symmetry.

**Case (b).** By Case (a) we may assume \( r > 0, s > 0 \) and \( j > 0 \) and by contradiction let’s assume that \( r \neq s \). If we look at Equations (6) and (7), we see that \( y \) is a square unless (i) \( s = 1 \) or (ii) both \( r = j = 1 \).

(i) Suppose \( s = 1 \). Since we’re done if \( r = s \) we may assume that \( r \geq 2, j \geq 2 \). From (8) we obtain

\[
(t^{2^j - 2^r} x^{2^r}) + (t^{1 - 2^j - 2^r} x^{2^j}) = t^{2^{s - 2r - 1}} + t^{1 - 2^{s - 1} - 2^j} + t + \frac{1}{t}
\]

or

\[
(t^{1 - 2^j - 2^r} x^{2^j - 1}) + (t^{1 - 2^{s - 1} - 2^j} x^{2^j - 1}) + t^{1 - 2^{s - 1} - 2^j} + t^{2^{s - 1} - 2^j - 1})^2 = t + \frac{1}{t}.
\]

Since \( j \geq 2 \) and \( r \geq 2 \) the left side is a square. The right side is not, contradiction.

(ii) Suppose \( r = j = 1 \). Again since we’re done if \( r = s \) we may assume \( s > 1 \). By (8) we have

\[
x^2 (t^{2^s - 1} + t^{2^2 - 1}) = t^{1 - 2^s} + t^{2^s - 1} + \frac{1}{t^{2^s - 1}} + t^{2^{s - 1}}.
\]
Let \( p \) be the simple zero of \( t \). Since \( 1 - 2^s < -2^{s-1} (s \geq 2) \), the right side has a pole of odd order at \( p \), while the left side is a square, so it only has poles of even order. This proves the theorem.

So the goal for the rest of this section is to show that the relations we used in the statement of the theorem are diophantine. To do that it will clearly be enough to show that the following four sets are diophantine:

\[
S := \{ t^{2^s} : s \in \mathbb{N} \}, \quad S' := \{ (t^{-1})^{2^s} : s \in \mathbb{N} \},
\]

\[
T := \left\{ (x, w) \in K^2 : \exists s \in \mathbb{N}, w = \left( \frac{x^2 + t^2 + t}{x^2 + t} \right)^{2^s} \right\},
\]

and

\[
T' := \left\{ (x, w) \in K^2 : \exists s \in \mathbb{N}, w = \left( \frac{x^2 + (t^{-1})^2 + t^{-1}}{x^2 + t^{-1}} \right)^{2^s} \right\}.
\]

It is enough to prove that \( S \) and \( T \) are diophantine, because we can replace \( t \) by \( t^{-1} \) and replace \( V \) by \( W := \left\{ \frac{1}{c} : c \in (V \setminus \{0\}) \right\} \cup \{0\} \) in Section 2. Then we can also replace \( t \) by \( t^{-1} \) and \( V \) by \( W \) in the whole proof to obtain diophantine definitions for \( S' \) and \( T' \).

Lemma 3.5 and Corollary 3.7 below will show that \( S \) is diophantine, and Corollary 3.11 will show that \( T \) is diophantine.

### 3.1. The set \( S = \{ t^{2^s} : s \in \mathbb{N} \} \) is diophantine.

To prove that \( S \) is diophantine, we first need a definition and a lemma:

**Definition 3.2.** Let \( w \in K \). The *height* of \( w \) is the degree of the zero divisor of \( w \).

**Remark.** Equivalently, we could have defined the height of \( w \in K \) to be the degree of the pole divisor of \( w \).

**Lemma 3.3.** Let \( w \in K \), let \( a, b \in C \). Then all the zeros of \( \frac{w+a}{w+b} \) are zeros of \( w+a \) and all the poles of \( \frac{w+a}{w+b} \) are zeros of \( w+b \). Furthermore, the height of \( \frac{w+a}{w+b} \) is equal to the height of \( w \).

**Proof.** This is Lemma 2.4 in \([Shl00]\). \(\square\)

**Lemma 3.4.** Let \( u, v, z \in \bar{K} := \hat{C}_KK \), assume that \( z \notin \hat{C}_K \), and let \( y \in \hat{C}_K(z) \). Assume that \( y, z \) do not have zeros or poles at any valuation of \( \bar{K} \) ramifying in the extension \( \bar{K}/\hat{C}_K(z) \) and that \( \bar{K}/\hat{C}_K(z) \) is separable. Moreover, assume

\[
y + z = u^4 + u \tag{9}
\]

\[
\frac{1}{y} + \frac{1}{z} = v^4 + v \tag{10}
\]

Then \( y = z^{4^k} \) for some \( k \geq 0 \).
Proof. Recall that for a field $F$ and a natural number $k$, $F^k = \{a^k : a \in F\}$. In $\tilde{C}_K(z)$ the zeros and poles of $z$ are simple. Assuming that $z$ satisfies the conditions of Lemma 3.4 thus amounts to assuming that all zeros and poles of $z$ are simple in $\tilde{K}$.

Equation (9) and the fact that $z$ has simple poles imply that $y \notin \tilde{C}_K$, so $y \in (\tilde{K})^4s$ only if $s \leq [\tilde{K} : \tilde{C}_K(z)]$. If $y = w^4$ with $w \in \tilde{C}_K(z)$, then $w + z = (u + w)^4 + (u + w)$ and $1/w + 1/z = (v + 1/w)^4 + (v + 1/w)$. So if we can prove that $w = z^{4s}$ for some $s \in \mathbb{N}$, then $y = w^4 = z^{4s+1}$. Hence we may assume without loss of generality that $y \notin (\tilde{C}_K(z))^4$.

Let $\frac{A}{B}$ be the divisor of $z$ in $\tilde{K}$, where $\mathcal{A}$ and $\mathcal{B}$ are relatively prime effective divisors. By assumption, all the prime factors of $\mathcal{A}$ and $\mathcal{B}$ are distinct. Also all the poles of $u^4 + u$ and $v^4 + v$ have orders divisible by 4.

Claim. The divisor of $y$ is of the form $\mathcal{E}^4\mathcal{D}$ where all the prime factors of $\mathcal{D}$ come from $\mathcal{A}$ or $\mathcal{B}$. Also the factors of $\mathcal{A}$ that appear in $\mathcal{D}$ will appear to the first power in $\mathcal{D}$ and the factors of $\mathcal{B}$ that appear in $\mathcal{D}$ occur to the power $-1$.

Proof of Claim. Let $t$ be a prime which is not a factor of $\mathcal{A}$ or $\mathcal{B}$. Without loss of generality assume $t$ is a pole of $y$. Then, since $\text{ord}_t z = 0$, we have

$$0 > \text{ord}_t y = \text{ord}_t (z + y) = \text{ord}_t (u^4 + u) \equiv 0 \mod 4.$$ 

Now let $t$ be a factor of $\mathcal{A}$ or $\mathcal{B}$. Again without loss of generality assume $t$ is a pole of $y$. If $t$ is a factor of $\mathcal{A}$, then $\text{ord}_t y = \text{ord}_t (y + z) = \text{ord}_t (u^4 + u)$. Hence $t$ is a pole of $u$, so $\text{ord}_t y \equiv 0 \mod 4$. If, however, $t$ is a factor of $\mathcal{B}$, there are two possibilities: Either $\text{ord}_t y = \text{ord}_t z = -1$ or again $\text{ord}_t y = \text{ord}_t (u^4 + u) \equiv 0 \mod 4$. This proves the claim.

On the other hand, $\mathcal{A}$ and $\mathcal{B}$ considered as divisors over $\tilde{C}_K(z)$ are prime divisors, and since $y \in \tilde{C}_K(z)$, we can deduce that the divisor of $y$ is of the form $\mathcal{E}^4\mathcal{A}^a\mathcal{B}^b$, with either, $a, b = 0$ or $a = 1, b = -1$, since the degree of the zero and the pole divisor must be the same.

Case I: $a = b = 0$.

Since no prime which is a zero of $y$ ramifies in the extension $\tilde{K}/\tilde{C}_K(z)$, the divisor of $y$ in $\tilde{C}_K(z)$ is also a fourth power of another divisor. In the rational function field $\tilde{C}_K(z)$ every degree 0 divisor is principal, so $y \in (\tilde{C}_K(z))^4$.

Case II: $a = 1, b = -1$.

In this case, the divisor of $\frac{y}{z}$ is of the form $\mathcal{E}^4$ and hence $\frac{y}{z} = f^4$ for some $f \in \tilde{C}_K(z)$ by the same argument as in Case I. Hence $y + z = u^4 + u$ can be rewritten as $z (\frac{y}{z} + 1) = z (f + 1)^4 = u^4 + u$. Since $f + 1$ is a rational function in $z$, we can rewrite this as

\begin{equation}
(11) \quad z \left(\frac{f_1}{f_2}\right)^4 = u^4 + u
\end{equation}
where $f_1, f_2$ are relatively prime polynomials in $\overline{C}_K[z]$, and $f_2$ is monic. Equation (11) shows: Any valuation which is a pole of $u$ is either a pole of $z$ or a zero of $f_2$. Let $c$ be a pole of $u$ which is a zero of $f_2$. Then, since $f_2$ is a polynomial in $z$, $c$ is not a pole of $z$. So we must have $\text{ord}_c f_2 = |\text{ord}_c u|$. Hence $s := f_2 \cdot u$ will have poles only at the valuations which are poles of $z$. Thus we can rewrite (11) in the form

$$z f_1^4 + s^4 = sf_2^3. \tag{12}$$

Furthermore, let $c$ be a zero of $f_2$. As pointed out above, $c$ is not a pole of $z$, so $c$ is not a pole of $s$. So we can deduce that for a zero $c$ of $f_2$ we have $\text{ord}_c (s^4 + z f_1^4) = \text{ord}_c (sf_2^3) \geq 3$. Thus $\text{ord}_c (d(s^4 + z f_1^4)) \geq 2$, so $\text{ord}_c (f_1^4 dz) \geq 2$. Here $dz$ denotes a Kähler differential. Since $c$ is unramified in the extension $\overline{K}/\overline{C}_K(z)$, $\text{ord}_c (dz) = 0$. Hence $\text{ord}_c (f_1^4) \geq 2$, i.e., $f_1$ has a zero at $c$. Since $f_1$ and $f_2$ are relatively prime polynomials, this implies that $f_2$ has no zeros, i.e., $f_2 = 1$. Hence $y$ is a polynomial in $z$. Exactly the same argument applied to $\frac{1}{y}$ shows that $\frac{1}{y}$ is a polynomial in $\frac{1}{z}$. Thus $y = z^l$ for some $l \geq 0$ and $y + z = z^l + z = u^4 + u$. If $y = z$, we are done. Otherwise this implies that all the poles of $y + z$ have order $l$ (the poles of $z$ are simple), and also, that all the poles of $y + z$ are divisible by 4. Hence $4|l$.

So in both cases, Case I and Case II, we could deduce that either $y = z$ or that $y \in \overline{C}_K(z)^4$. Since we assumed that $y \notin \overline{C}_K(z)^4$ this concludes the proof of the Claim. \qed

**Lemma 3.5.** For all $c, c' \in V$ let $t_{c,c'} := \frac{t+c}{t-c}$. Let $w, v, u, u_{d,d'}, v_{d,d'}$ be elements of $K$ such that $\forall c \in V \exists d \in V_c$ such that $\forall c' \in V \exists d' \in V_{c'}$ such that the following equations are satisfied:

$$w + t = u^4 + u \tag{13}$$
$$\frac{1}{w} + \frac{1}{t} = v^4 + v \tag{14}$$
$$w_{d,d'} = \frac{w + d}{w + d'} \tag{15}$$
$$w_{d,d'} + t_{c,c'} = u_{d,d'} + u_{d,d'} \tag{16}$$
$$\frac{1}{w_{d,d'}} + \frac{1}{t_{c,c'}} = v_{d,d'} + v_{d,d'}. \tag{17}$$

Then $w = t^4s$ for some natural number $s$.

**Proof.** Recall that the divisor of $t$ in $K$ is of the form $p/q$, and that $p$ and $\Omega$ are the primes of $C_K(t)$ lying below $p$ and $q$, respectively. Thus the degree of $\Omega$ is one. Similarly, for all $c \in V$ the degree of the primes $p_c$ in $C_K(t)$ is one. Hence $\Omega$ and all the $p_c$’s will remain prime in the constant field extension $\overline{C}_K(t)/C_K(t)$. By Lemma 6.16 in [Sh100] their factors will be unramified in
the extension $\bar{K}/\bar{C}_K(t)$. Hence for all $c,c' \in V$, $t_{c,c'}$ has neither zeros nor poles at any prime ramifying in the extension $\bar{K}/\bar{C}_K(t)$.

In the second paragraph of the proof of Lemma 2.6 of [Shl00], pp. 471-472, translated to our notation, Shlapentokh proves that for some $c_0 \in V$ there exists a subset $V'$ of $V$ containing $n + 1$ elements, not containing $c_0$, and such that for any $d_0 \in V_{c_0}$, for all $c' \in V'$, for any $d' \in V_{c'}$, $w_{d_0,d'}$ does not have zeros or poles at any prime ramifying in the extension $\bar{K}/\bar{C}_K(t)$. Her argument uses the fact that there are exactly $r$ primes ramifying in the extension $\bar{K}/\bar{C}_K(t)$, and it does not use the characteristic of $K$, so the same proof works here. We have two cases to consider:

Case I: $w \in C_K(t)$.

If $w$ is in $C_K(t)$, then pick a $d_0 \in V_{c_0}$ and for some $c' \in V'$ pick a $d' \in V_{c'}$ such that (16) and (17) are satisfied. Then $w_{d_0d'} \in C_K(t)$, and we can apply Lemma 3.4 to $t_{c_0,c'}$ instead of $t$, and to $w_{d_0d'}$ to conclude that $w_{d_0d'} = t^{s_0}_{c_0,c'}$ for some $s \geq 0$. (Note that $C_K(t) = C_K(t_{c_0,c'})$.) So

$$\frac{w + d_0}{w + d'} = (t_{c_0,c'})^{s_0}.\)

If $s = 0$, then we can check that $w = t$. (See the last part of Lemma 3.3.) Otherwise write $1 + \frac{d_0 + d'}{w + d'} = (t_{c_0,c'})^{s_0}$.

Hence $w + d' = \left(\frac{1}{t_{c_0,c'} + 1}\right)^{s_0} \cdot (d_0 + d')$. Since $(d_0 + d')$ is an element of $C$ and hence a fourth power this implies that $w + d' \in (C_K(t))^4$, and hence $w \in (C_K(t))^4$, say $w = \bar{w}^4$. We can rewrite Equations (13) and (14) as

\begin{align*}
\bar{w} + t &= (u + \bar{w})^4 + (u + \bar{w}) \\
\frac{1}{w} + \frac{1}{t} &= \left(\frac{v + \frac{1}{w}}{w}\right)^4 + \left(\frac{v + \frac{1}{w}}{w}\right).
\end{align*}

(18) (19)

Also

$$w_{d,d'} = \frac{w + d}{w + d'} = \frac{w + \tilde{d}^4}{w + d^4} = \left(\frac{\bar{w} + \bar{d}}{\bar{w} + d'}\right)^4$$

for $d \in V_c, d' \in V_{c'}$ and some suitable $\tilde{d} \in V_c$ and $\tilde{d'} \in V_{c'}$. This lets us rewrite Equations (16) and (17) in a similar fashion. So we can rewrite Equations (13) through (17), and $\bar{w} \in C_K(t)$. Equation (13) and the fact that $t$ has only simple zeros imply that $w \notin C_K$. Hence after finitely many iterations we must be in the position where $s = 0$.

Case II: $w \notin C_K(t)$.

In this case we will derive a contradiction. $w \notin C_K(t)$ would imply that $w_{d,d'} \notin C_K(t)$ for all $d$ and $d'$. 


By putting $\alpha := u^2 + u$ we can rewrite Equation (13) as

$$w + t = \alpha^2 + \alpha.$$  

Similarly by putting $\beta := v^2 + v$, $\alpha_{d,d'} := u_{d,d'}^2 + u_{d,d'}$, $\beta_{d,d'} := v_{d,d'}^2 + v_{d,d'}$
we can rewrite Equations (14), (16) and (17) as

$$\frac{1}{w} + \frac{1}{t} = \beta^2 + \beta,$$
$$w_{d,d'} + t_{c,c'} = \alpha_{d,d'}^2 + \alpha_{d,d'},$$
$$\frac{1}{w_{d,d'}} + \frac{1}{t_{c,c'}} = \beta_{d,d'}^2 + \beta_{d,d'}.$$

Let $c_0 \in V$ be as above. By the same argument as in [Shl00], p. 472, with $p$ replaced by 2, (20) through (23) imply that $\exists d_0 \in V_{c_0}$ such that $\forall c' \in V', \exists d' \in V_d$ such that the divisor of $w_{d_0,d'}$ is of the form $A^2 p_d^a p_{d_0}^b$. Here $p_d$ and $p_{d'}$ are prime divisors, $a$ is either $-1$ or 0, and $b$ is either 1 or 0. Now the proof follows word for word that of Lemma 2.6 in [Shl00], p. 472 with $p$ replaced by 2 to prove that in this case $w = \bar{w}^2$ with $\bar{w} \in K$. For this part of the proof we only used Equations (20) through (23). Now we can rewrite Equations (20) through (23) with $w$ replaced by $\bar{w}$. Since $w \notin C_K(t), \bar{w} \notin C_K(t)$. So we can keep replacing $w$ by its square root over and over, contradicting that $w \in K^{2\pi}$ only if $s \leq [K : C_K(t)]$. So $w = t^{4s}$ for some $s \in \mathbb{N}$.

\begin{corollary}
The set $S_1 := \{t^{4s} : s \in \mathbb{N}\}$ is diophantine over $K$.
\end{corollary}

\begin{proof}
Lemma 3.5 shows that an element $w \in K$ satisfying Equations (13) through (17) must be of the form $w = t^{4s}$ for some $s \in \mathbb{N}$. What we have left to show is that if $w = t^{4s}$ for some $s \in \mathbb{N}$, then we can satisfy Equations (13) through (17). If $w = t$, let $u = 0$ and $v = 0$. For the general case we use the fact that for any $x \in K$ and any $s \in \mathbb{N}$ we have

$$x^{4s} - x = (x^{4s-1} + x^{4s-2} + \cdots + x)^4 - (x^{4s-1} + x^{4s-2} + \cdots + x).$$

So if $w = t^{4s}$ with $s \geq 1$, let $u = t^{4s-1} + \cdots + t^4 + t$. For $v$ take

$$\frac{1}{t^{4s-1}} + \cdots + \frac{1}{t^4} + \frac{1}{t}.$$ 

Now fix $c \in V$. To satisfy the other equations we can use the same argument, if we can show that $\exists d \in V_c$ such that $\forall c' \in V \exists d' \in V_d$ such that $w_{d,d'} = (t_{c,c'})^{4s}$. This is done in Corollary 2.7 in [Shl00].

\begin{corollary}
The set $S = \{t^{2s} : s \in \mathbb{N}\}$ is diophantine over $K$.
\end{corollary}

\begin{proof}
This follows from the fact that

$$w \in S \iff (w \in S_1 \text{ or } \exists z \in K (z^2 = w \text{ and } z \in S_1)).$$

\end{proof}
3.2. The set \( T = \left\{ (x, w) \in K^2 : \exists s \in \mathbb{N}, w = \left( \frac{x^2 + t^2 + t}{x^2 + t} \right)^{2^s} \right\} \) is diophantine over \( K \).

**Lemma 3.8.** Let \( x \in K \). Let \( t \) be as above, i.e., \( \overline{K}/\overline{C}_K(t) \) is separable and the divisor of \( t \) is of the form \( \frac{p}{q} \). Let \( u = \frac{x^2 + t^2 + t}{x^2 + t} \), and let \( a \in C, a \neq 1 \). Then \( u + a \) has only simple zeros and simple poles, except possibly for zeros at \( p, q \) or primes ramifying in the extension \( \overline{K}/\overline{C}_K(t) \).

**Proof.** First we will show that the zeros of \( u + a \) away from the ramified primes and \( p \) and \( q \) are simple. By Lemma 4.4 in [Shi96] it is enough to show that \( u + a \) and \( \frac{du}{dt} \) do not have common zeros. We have

\[
\frac{d(u + a)}{dt} = \frac{(x^2 + t^2 + t) + (x^2 + t)}{(x^2 + t)^2} = \frac{t^2}{(x^2 + t)^2} \quad \text{and} \quad u + a = \frac{x^2 + t^2 + t}{x^2 + t} + a = 1 + a + \frac{t^2}{x^2 + t}.
\]

Suppose \( c \) is a zero of \( d(u + a)/dt \) satisfying the above conditions. Then \( c \) is not a zero of \( t \), so \( c \) must be a pole of \( x^2 + t \), i.e., a pole of \( x \). If \( c \) is a pole of \( x \), then it is a zero of \( \frac{t^2}{x^2 + t} \), and hence not a zero of \( 1 + a + \frac{t^2}{x^2 + t} \). Hence \( d(u + a)/dt \) and \( u + a \) have no zeros in common, except possibly the ones mentioned above.

We will now show that all poles at above described valuations are simple: Since \( u \) and \( u + a \) have the same poles, it is enough to show that the poles of \( u \) are simple. \( u \) has simple poles if and only if the zeros of \( u^{-1} \) are simple. So we’ll show that the zeros of \( v = u^{-1} \) are simple by doing exactly the same thing as above. Let \( v := u^{-1} = \frac{x^2 + t}{x^2 + t^2 + t} \). Then

\[
\frac{dv}{dt} = \frac{(x^2 + t + x^2 + t^2 + t)}{(x^2 + t^2 + t)^2} = \frac{t^2}{(x^2 + t^2 + t)^2} \quad \text{and} \quad v = \frac{x^2 + t}{x^2 + t^2 + t} = 1 - \frac{t^2}{(x^2 + t^2 + t)^2}.
\]

Again let \( c \) be a zero of \( dv/dt \) satisfying the above conditions. Again \( c \) has to be a pole of \( x \). So \( c \) is a zero of \( v \), but not a zero of \( 1 - \frac{t^2}{x^2 + t^2 + t} \), since \( c \) is not a zero or pole of \( t \). Hence all the zeros of \( u^{-1} \) are simple except possibly for the ones mentioned above. \( \square \)

**Lemma 3.9.** Let \( x, v \in K^* \), let \( u := \frac{x^2 + t^2 + t}{x^2 + t} \). For all \( c, c' \in V \), \( g \in \{-1, 1\} \) let

\[
u_{c, c', g} := \frac{u^g + c}{u^g + c'}.
\]
For \(d \in V_c, d' \in V_c', g \in \{-1, 1\}\) let
\[
v_{d,d',g} := \frac{v^g + d}{v^g + d'}.
\]

In addition assume that \(\forall c \in V \exists d \in V_c\) such that \(\forall c' \in V \exists d' \in V_c'\) such
\[
\text{that the following equations hold for for } e, g \in \{-1, 1\}, \text{ and some } s \in \mathbb{N}:
\]
\begin{align*}
&u_{d,d',g}^c + u_{c,c',g}^e = \sigma_{d,d',e,g}^4 + \sigma_{d,d',e,g} \quad \text{for all } c, c', g, e, \quad \text{(24)} \\
&v_{d,d',g}^{2c} t^{4v} + u_{c,c',g}^{2e} t = \lambda_{d,d',e,g}^4 + \lambda_{d,d',e,g} \quad \text{for all } c, c', g, e, \quad \text{(25)} \\
&(u^g + c)^e + (v^g + d)^e = \mu_{d,e,g}^4 + \mu_{d,e,g} \quad \text{for all } c, c', g, e, \quad \text{(26)}
\end{align*}

Then for some natural number \(m\), \(v = u^{4m}\).

**Proof.** It is sufficient to prove that the result is valid in \(\tilde{K} := \tilde{C}_K K\), so we
work in \(K\). We will first prove the following:

**Claim.** For all \(c, c' \in V, g \in \{-1, 1\}, \ u_{c,c',g}\) has no multiple zeros or poles
except possibly at the primes ramifying in \(\tilde{K}/\tilde{C}_K(t)\) or \(p\) or \(q\).

**Proof of Claim.** By Lemma 3.3 we have that for all \(c, c', g\) as above all the
poles of \(u_{c,c',g}\) are zeros of \(u^g + c\) and all the zeros of \(u_{c,c',g}\) are zeros of
\(u^g + c\). By Lemma 3.8 and by assumption on \(c\) and \(c'\), all the zeros of \(u^g + c\)
and \(u^g + c\) are simple, except possibly for zeros at \(p, q\) or primes ramifying
in the extension \(\tilde{K}/\tilde{C}_K(t)\). This proves the claim.

Also since \(\frac{du}{dt} \neq 0\), \(u\) is not a second power in \(\tilde{K}\). We will show the
following: (a) If \(s = 0\), then \(u = v\), and (b) if \(s > 0\), then \(v\) is a fourth power
of some element in \(\tilde{K}\).

**Case** (a): Suppose that \(s = 0\), and set \(g = 1\).

Again, using Shlapentokh’s argument in Lemma 2.6 of [Shl00] there exists
\(c_0 \in V\) and \(V' \subseteq (V - \{c_0\})\) containing \(n + 1\) elements, such that for all
\(d_0 \in V_{c_0}\), for all \(c' \in V'\), and for all \(d' \in V_{c'}\), \(u_{c_0,c',1}\) and \(v_{d_0,d',1}\) have no zeros or poles at the primes of \(\tilde{K}\) ramifying in the extension \(\tilde{K}/\tilde{C}(t)\) or at \(p\) or \(q\).

For indices selected in this way, all the poles and zeros of \(u_{c_0,c',1}\) are simple.
Pick \(d_0 \in V_{c_0}\) and for all \(c' \in V'\) pick \(d' \in V_{c'}\) such that Equations (24) and
(25) are satisfied. Equations (25) and (24) imply:
\begin{align*}
&v_{d_0,d',1}^2 t^{4v} + u_{c_0,c',1}^2 t = \lambda_{d_0,d',1,1}^4 + \lambda_{d_0,d',1,1} \quad \text{for all } c, c', g, e, \quad \text{(27)} \\
&v_{d_0,d',1}^2 + u_{c_0,c',1}^2 = \sigma_{d_0,d',1,1}^8 + \sigma_{d_0,d',1,1}^2 \quad \text{for all } c, c', g, e, \quad \text{(28)}
\end{align*}

From (27) and (28) we obtain (since \(s = 0\))
\[
\lambda_{d_0,d',1,1}^4 + \lambda_{d_0,d',1,1} = t(\sigma_{d_0,d',1,1}^8 + \sigma_{d_0,d',1,1}^2).
\]

All the poles of \(\lambda_{d_0,d',1,1}\) and \(\sigma_{d_0,d',1,1}\) are poles of \(u_{c_0,c',1}\), \(v_{d_0,d',1}\) or \(t\), and
thus are not at any valuation ramifying in the extension \(\tilde{K}/\tilde{C}_K(t)\). By
Lemma 4.4 applied to $\sigma = \sigma_{d_0,d',1}^2$ and (28)

$$v_{d_0,d',1}^2 + u_{c_0,c',1}^2 = 0.$$  

Thus $v_{d_0,d',1} = u_{c_0,c',1}$. From here on the proof is word for word like the proof of Lemma 2.10 in [Shi00], top of p. 477, showing that if $s = 0$, then $u = v$.

Case (b): Suppose now that $s > 0$. Again set $g = 1$. Let $c_0$ and $V'$ be selected as above.

Again we can pick $d_0 \in V_{c_0}$ and for all $c' \in V'$ we can pick $d' \in V_{c'}$ such that Equations (24) through (26) are satisfied and such that the corresponding $u_{c_0,c',1}$ and $v_{d_0,d',1}$ do not have zeros or poles at the primes of $\bar{K}$ ramifying in the extension $\bar{K}/\bar{C}(t)$ or at $p$ or $q$. We can use the same argument as in Lemma 3.4 to show that either:

(i) For all $d'$ chosen as above the divisor of $v_{d_0,d',1}$ in $\bar{K}$ is a fourth power of another divisor, or

(ii) for some $c' \in V'$ and some $d' \in V_{c'}$ and some prime $t$ not ramifying in $\bar{K}/\bar{C}(t)$ and not equal to $p$ or $q$, $\text{ord}_{t} v_{d_0,d'} \in \{ 1, -1 \}$.

In Case (i), because of our choice of the $v_{d_0,d'}$’s and Proposition 4.3, a short calculation shows that $v \in \bar{K}^4$:

$$v_{d_0,d',1}^{-1} = \frac{v + d_0}{v + d'} = \frac{1 + (\bar{d'} + \bar{d})}{\frac{1}{v + d_0}} = (\bar{d'} + \bar{d}) \left( \frac{1}{d' + d_0} + \frac{1}{v + d_0} \right),$$

where $d_0 \in V_{c_0}$ is fixed, and we have $d' \in V_{c'}$, and all $d'$ are distinct. Also $V'$ contains $n + 1$ elements, so by Proposition 4.3 applied to $\frac{1}{v + d'}$ we have that $\frac{1}{v + d'} \in \bar{K}^4$. This implies that $v \in \bar{K}^4$. This finishes Case (i).

So assume now that we are in Case (ii): Without loss of generality, assume that $t$ is a pole of $v_{d_0,d'}$ (and hence neither a zero nor a pole of $t$).

Again look at Equations (27) and (28). Since $t$ does not have a pole or a zero at $t$ and since the right hand sides of Equations (27) and (28) only have poles of order $\geq 4$,

$$\text{ord}_t (v_{d_0,d',1}^2 t^4 + u_{c_0,c',1}^2 t) = \text{ord}_t (\lambda_{d_0,d',1}^4 + \lambda_{d_0,d',1,1}) \geq 0$$  

and

$$\text{ord}_t (v_{d_0,d',1}^2 + u_{c_0,c',1}^2) = \text{ord}_t (\sigma_{d_0,d',1}^8 + \sigma_{d_0,d',1,1}^2) \geq 0.$$  

Thus

$$\text{ord}_t v_{d_0,d',1}^2 (t^4 + t) \geq 0.$$  

Hence it follows that $\text{ord}_t (t^4 + t) \geq 2|\text{ord}_t v_{d_0,d',1}|$. But in $\bar{C}_K(t)$ all the zeros of $t^4 + t$ are simple. So this function can have multiple zeros only at primes ramifying in the extension $\bar{K}/\bar{C}_K(t)$. But by assumption $t$ is not one
of these primes, and so we have a contradiction unless \( v \in \overline{K}^4 \). This shows that if \( s > 0 \), then Equations (24) through (26) can be rewritten in the same fashion as in Lemma 3.5 with \( v \) replaced by its fourth roots, and in (25) \( s \) is replaced by \( s - 1 \). Therefore, after finitely many iterations of this rewriting procedure we will be in the case of \( s = 0 \), which was treated in Case (a).

Hence, for some natural number \( m \), \( v = u^{4^m} \).

**Corollary 3.10.** The set \( T_1 := \{ (x, w) \in K^2 : \exists s \in \mathbb{N}, w = \left( \frac{x^2 + t^2 + t}{x^2 + t} \right)^{4^s} \} \) is diophantine over \( K \).

**Proof.** Let \( x \in K \), and let \( u = \frac{x^2 + t^2 + t}{x^2 + t} \). Lemma 3.9 shows that an element \( v \in K \) satisfying Equations (24) through (26) must be of the form \( v = u^{4^k} \) for some \( k \in \mathbb{N} \). So we have to show now that if \( v = u^{4^k} \) for some \( k \in \mathbb{N} \), then Equations (24) through (26) can be satisfied. The proof of this is almost identical to Corollary 3.6. \( \square \)

**Corollary 3.11.** The set \( T := \{ (x, w) \in K^2 : \exists s \in \mathbb{N}, w = \left( \frac{x^2 + t^2 + t}{x^2 + t} \right)^{2^s} \} \) is diophantine over \( K \).

**Proof.** This follows from the fact that \( (x, w) \in T \iff (x, w) \in T_1 \) or \( \exists z \in K (z^2 = w \text{ and } (x, z) \in T_1) \).

**Theorem 3.12.** The set \( \{ (x, y) \in K^2 : \exists s \in \mathbb{N}, y = x^{2^s} \} \) is diophantine over \( K \).

**Proof.** By Corollary 3.7, Corollary 3.11 and the remark after Theorem 3.1, the sets \( S, S', T \), and \( T' \) are diophantine. Together with Theorem 3.1 this finishes the proof. \( \square \)

### 4. Appendix.

In the appendix we give proofs for Proposition 4.3 and Lemma 4.4. Both were used in Lemma 3.9.

**Lemma 4.1.** Let \( F/G \) be a finite extension of fields of positive characteristic \( p \). Let \( \alpha \in F \) be such that all the coefficients of its monic irreducible polynomial over \( G \) are in \( G^{p^2} \). Then \( \alpha \in F^{p^2} \).

**Proof.** This is Lemma 2.1 in [Shl00] with \( p \) replaced by \( p^2 \), and the same proof works here. \( \square \)

**Corollary 4.2.** Let \( F/G \) be a finite separable extension of fields of positive characteristic \( p \). Let \( [F : G] = n \). Let \( x \in F \) be such that \( F = G(x) \), and such that for distinct \( a_0, \ldots, a_n \in G \), \( N_{F/G}(a_i^{p^2} - x) = y_i^{p^2} \) with \( y_i \in G \). Then \( x \in F^{p^2} \).
Proof. This is very similar to Lemma 2.2 in [Shl00]. Let \(H(T) = A_0 + A_1 T + \cdots + T^n\) be the irreducible polynomial of \(x\) over \(G\). Then \(H(a_i^p) = y_i^p\) for \(i \in \{0, \ldots, n\}\). This gives us the following linear system of equations:

\[
\begin{pmatrix}
1 & a_0^p & \cdots & a_0^{p(n-1)} & a_0^{p^2 n} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
1 & a_n^p & \cdots & a_n^{p(n-1)} & a_n^{p^2 n}
\end{pmatrix}
\begin{pmatrix}
A_0 \\
\vdots \\
1
\end{pmatrix}
= \begin{pmatrix}
y_0^p \\
\vdots \\
y_n^p
\end{pmatrix}.
\]

We can use Cramer’s rule to solve the system and to conclude that \(A_i \in G^p\) for all \(i\). Then by the previous lemma, \(x \in F^p\). \(\Box\)

Now we can apply the corollary to our situation:

**Proposition 4.3.** Let \(v \in \widetilde{K}\), and assume that for some distinct \(a_0, \ldots, a_n\) in \(C\), the divisor of \(v + a_i\) is of the form \(D_i^p\) for divisors \(D_i\) of \(\widetilde{K}\), \(i = 0, \ldots, n\). Moreover, assume that for all \(i\), \(v + a_i\) does not have zeros or poles at any prime ramifying in the extension \(\widetilde{K}/\widetilde{C}_K(t)\). Then \(v \in \widetilde{K}^p\).

**Proof.** This is almost the same as Lemma 2.9 in [Shl00]: First assume that \(v \in \widetilde{C}_K(t)\). Since \(v + a_i\) does not have any zeros or poles at primes ramifying in the extension \(\widetilde{K}/\widetilde{C}_K(t)\), the divisor of \(v + a_i\) in \(\widetilde{C}_K(t)\) is of the form \(E_i^p\).

In \(\widetilde{C}_K(t)\) every divisor of degree zero is principal, so \(v + a_i \in (\widetilde{C}_K(t))^p\) and hence \(v \in (\widetilde{C}_K(t))^p\). Therefore \(v \in \widetilde{K}^p\).

So now assume that \(v \notin \widetilde{C}_K(t)\). From our assumption on \(v + a_i\) it follows that in \(\widetilde{C}_K(t,v)\) the divisor of \(v + a_i\) is a \(p\) power of another divisor. Since the divisor of \(N_{\widetilde{C}_K(t,v)/\widetilde{C}_K(t)}(v + a_i)\) is equal to the corresponding norm of the divisor of \((v + a_i)\), it follows that the divisor of the \(\widetilde{C}_K(t,v)/\widetilde{C}_K(t)\) norm of \((v + a_i)\) is of the form \(N_i^p\), and hence \(N_{\widetilde{C}_K(t,v)/\widetilde{C}_K(t)}(v + a_i) \in (\widetilde{C}_K(t))^p\).

Now apply Corollary 4.2 with \(G = \widetilde{C}_K(t)\) and \(F = \widetilde{C}_K(t,v)\). \(\Box\)

**Lemma 4.4.** Let \(\sigma, \mu \in K\). Assume that all the primes that are poles of \(\sigma\) or \(\mu\) do not ramify in the extension \(\widetilde{K}/\widetilde{C}_K(t)\). Moreover assume that

\[
(31) \quad t(\sigma^4 + \sigma) = \mu^4 + \mu.
\]

Then \(\sigma^4 + \sigma = \mu^4 + \mu = 0\).

**Proof.** Let \(A, B\) be effective divisors of \(K\), relatively prime to each other and to \(p\) and \(q\), such that the divisor of \(\sigma\) is of the form \(\frac{A}{B} p^i q^k\), where \(i\) and \(k\) are integers.

**Claim 1.** For some effective divisor \(C\) relatively prime to \(B, p\) and \(q\), some integers \(j, m\), the divisor of \(\mu\) is of the form \(\frac{C}{B} p^j q^m\).
Proof of Claim 1. Let \( t \) be a pole of \( \mu \) such that \( t \neq p \) and \( t \neq q \). Then
\[
0 > 4 \operatorname{ord}_t \mu = \operatorname{ord}_t (\mu^4 + \mu) = \operatorname{ord}_t (t(\sigma^4 + \sigma)) = \operatorname{ord}_t (\sigma^4 + \sigma) = 4 \operatorname{ord}_t \sigma.
\]
Conversely, let \( t \) be a pole of \( \sigma \) such that \( t \neq p \) and \( t \neq q \). Then
\[
0 > 4 \operatorname{ord}_t \sigma = \operatorname{ord}_t (\sigma^4 + \sigma) = \operatorname{ord}_t (t(\sigma^4 + \sigma)) = \operatorname{ord}_t (\mu^4 + \mu) = 4 \operatorname{ord}_t \mu.
\]
This proves the claim.

By the Strong Approximation Theorem there exists \( b \in K^* \) such that the divisor of \( b \) is of the form \( \frac{BD}{q} \), where \( D \) is an effective divisor relatively prime to \( A, C, p \) and \( q \) and \( i \) is a natural number.

Claim 2.

\[
bs = st^i \quad \text{and} \quad b\mu = s_2t^j,
\]
where \( s_1, s_2 \) are integral over \( C_K[t] \) and have zero divisors relatively prime to \( p \) and \( B \).

Proof of Claim 2. The divisor of \( bs \) is
\[
\frac{BD}{q^i} A p^i q^k = DA p^i q^{k-l} = (DA q^{k-l+i}) \left( \frac{p^i}{q^i} \right).
\]
Thus the divisor of \( s_1 := bs/t^i \) is of the form \( DA q^{k-l+i} \). Therefore \( q \) is the only pole of \( s_1 \), so \( s_1 \) is integral over \( C_K[t] \). By construction \( A \) and \( D \) are relatively prime to \( p \) and \( B \). A similar argument applies to \( s_2 := b\mu/t^j \). This proves the claim.

Multiplying (31) by \( b^4 \) we obtain the following equation (using the claim):
(32)
\[
t(s_1^4 t^{4i} + b^3 s_1 t^4) = s_2^4 t^{4j} + b^3 s_2 t^j.
\]
Suppose \( i < 0 \). Then the left side of (32) has a pole of order \(|4i + 1| \) at \( p \). This would imply that \( j < 0 \), and the right side has a pole of order \(|4j| \) at \( p \), contradiction. Thus we can assume that \( i, j \) are both nonnegative. We can rewrite (32) as
\[
(s_1^4 t^{4i+1} + s_2^4 t^{4j}) = b^3(s_1 t^{i+1} + s_2 t^j).
\]
Let \( t \) be any prime factor of \( B \) in \( \bar{K} \). Then \( t \) does not ramify in the extension \( \bar{K}/C_K(t) \) by our assumption on \( \sigma \). Also \( t \) is not a pole of \( s_1, s_2 \) or \( t \).

Thus
\[
\operatorname{ord}_t (s_1^4 t^{4i+1} + s_2^4 t^{4j}) = \operatorname{ord}_t (b^3(s_1 t^{i+1} + s_2 t^j)) \geq 3.
\]
We have
\[
0 < \operatorname{ord}_t (d(s_1^4 t^{4i+1} + s_2^4 t^{4j})) = \operatorname{ord}_t (s_1^4 d(t^{4i+1})) = \operatorname{ord}_t (s_1^4) + \operatorname{ord}_t (d(t^{4i+1})).
\]
Since \( t \) is unramified in the extension \( \bar{K}/C_K(t) \) and since \( t \) is not a zero or a pole of \( t \), \( \operatorname{ord}_t (d(t^{4i+1})) = 0 \). So \( s_1 \) has a zero at \( t \). This, however, is impossible, because \( t \) is a prime factor of \( B \), but the zero divisor of \( s_1 \) is
relatively prime to $B$. So $B$ must be the trivial divisor. This implies that in (31) all the functions are integral over $C_K[t]$, i.e., they can have poles at $q$ only. So if $\mu$ is not constant, it must have a pole at $q$. But then the left side of (31) has a pole at $q$ of odd order, while the right side of (31) has a pole at $q$ of even order, which is a contradiction.

Thus $\mu$ must be a constant. But if a function $h \in K$ is integral over $C_K[t]$, and $t \cdot h$ is constant, then $h = 0$. Thus $\sigma^4 + \sigma = 0$. Then $\mu^4 + \mu = 0$ also. □

Remark. A. Shlapentokh informed the author by email on March 31, 2003, that she has an argument in her paper [Compositio Math., 132 (2002), pp. 99-120] that reduces the case of finite transcendence degree to transcendence degree 1. Together with the result in this paper this implies that Hilbert’s Tenth Problem is undecidable for function fields $F$ of characteristic 2, finitely generated over a field $C$ that is algebraic over a finite field, and such that $C$ has an extension of degree 2.

References


Received July 1, 2002 and revised September 11, 2002.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CA 94720
E-mail address: eisen@math.berkeley.edu
COMPLEXITY AND HEEGAARD GENUS OF
AN INFINITE CLASS OF COMPACT 3-MANIFOLDS

ROBERTO FRIGERIO, BRUNO MARTELLI, AND CARLO PETRONIO

Using the theory of hyperbolic manifolds with totally geodesic boundary, we provide for every \( n \geq 2 \) a class \( \mathcal{M}_n \) of such manifolds all having Matveev complexity equal to \( n \) and Heegaard genus equal to \( n + 1 \). All the elements of \( \mathcal{M}_n \) have a single boundary component of genus \( n \), and \( \# \mathcal{M}_n \) grows at least exponentially with \( n \).

This paper is devoted to the investigation of the class \( \mathcal{M}_n \) of orientable compact 3-manifolds having an ideal triangulation with \( n \geq 2 \) tetrahedra and a single edge. We show in particular for each \( M \) in \( \mathcal{M}_n \) that the Heegaard genus of \( M \) is equal to \( n + 1 \), and that the complexity of \( M \) in the sense of Matveev is equal to \( n \). Moreover we prove that the classical invariants, such as homology and those of Turaev and Viro, cannot distinguish two different members of \( \mathcal{M}_n \) from each other. However, using the fact that each \( M \) in \( \mathcal{M}_n \) carries a hyperbolic metric with totally geodesic boundary, we prove that \( M \) has a unique ideal triangulation with \( n \) tetrahedra. We exploit this property showing that the number of elements of \( \mathcal{M}_n \) grows at least exponentially with \( n \). This implies in particular the previously unknown fact that the rate of growth of the number of orientable boundary-irreducible acylindrical manifolds of complexity \( n \) is also at least exponential in \( n \). The class \( \mathcal{M}_n \) was already considered in [8], but none of our results was covered there.

1. Manifolds with a one-edged triangulation.

In this section we introduce the class of manifolds we are interested in, and we prove their many remarkable topological and geometric properties.

Ideal triangulations and spines. We begin by recalling some definitions. An ideal tetrahedron is a tetrahedron with its vertices removed. An ideal triangulation of a compact 3-manifold \( M \) with boundary is a realization of the interior of \( M \) as a gluing of some ideal tetrahedra, induced by a simplicial pairing of the faces. A spine of \( M \) is a compact polyhedron \( P \) such that \( M \setminus P = \partial M \times [0, 1) \). A 2-dimensional polyhedron \( Q \) is quasi-standard if every point has a neighbourhood homeomorphic to one of the polyhedra shown in Fig. 1.
We denote by $V(Q)$ the set of points having regular neighbourhoods of type (3), and by $S(Q)$ the set of points having regular neighbourhoods of type (2) or (3). If $Q \setminus S(Q)$ consists of open cells and $S(Q) \setminus V(Q)$ consists of open edges, we say that $Q$ is \textit{standard}, and we call \textit{faces} the components of $Q \setminus S(Q)$. An ideal triangulation of $M$ defines in a natural way a dual standard polyhedron, which is in fact a spine of $M$ (see Fig. 2).

We now define the class of manifolds investigated in this paper. For every integer $n \geq 2$ we set

$$\mathcal{M}_n = \{ M : \dim(M) = 3, \ M \text{ is compact and orientable, } \partial M \neq \emptyset, \ M \text{ admits an ideal triangulation with one edge and } n \text{ tetrahedra} \}.$$ 

The class $\mathcal{M}_n$ can be defined in various equivalent ways, as the next lemma shows.

\textbf{Lemma 1.1.} Let $M$ be a compact 3-manifold and let $T$ be an ideal triangulation of $M$ consisting of $n$ tetrahedra. The following facts are equivalent:

1. $T$ has one edge;
2. $\chi(M) = 1 - n$;
3. $\chi(M) = 1 - n$ and $\partial M$ is connected.
Proof. Set $\hat{M} = M/\partial M$, let $|\partial M|$ be the number of components of $\partial M$, and let $x$ be the number of edges of $T$. We have $\chi(\hat{M}) = \chi(M) - \chi(\partial M) + |\partial M|$ and $\chi(\partial M) = 2\chi(M)$. Extending $T$ to a cellularization of $\hat{M}$, we also get

$$\chi(\hat{M}) = |\partial M| - x + 2n - n = |\partial M| - x + n.$$ 

Summing up, we get $\chi(M) = x - n$, which shows $\text{(1)} \Leftrightarrow \text{(2)}$.

We are left to show that if $x = 1$ then $|\partial M| \leq 2$. Since there is a single edge, we have $|\partial M| \leq 2$. Suppose $A$ and $B$ are distinct components of $\partial M$, and examine a triangular face $F$ of $T$. Let $e_1$, $e_2$, and $e_3$ be the edges of $F$ viewed abstractly (i.e., before the embedding in $M$). Since $e_1, e_2, e_3$ become the same edge in $M$, they should all join $A$ to $B$, which is clearly impossible: If $e_1$ and $e_2$ join $A$ to $B$, then $e_3$ joins either $A$ or $B$ to itself. □

**Topological and geometric properties.** Before proving our main theorem, we recall the definition of Matveev complexity of a compact 3-manifold with boundary, and the notion of hyperbolic 3-manifold with geodesic boundary. A compact 2-dimensional polyhedron $Q$ is said to be *simple* if the link of every point in $Q$ is contained in the 1-skeleton $K$ of the tetrahedron. (Note that a standard polyhedron is obviously simple.) A point having the whole of $K$ as a link is called a *vertex*, and its regular neighbourhood is as shown in Fig. 1-(3). This implies that the set $V(Q)$ of the vertices of $Q$ consists of isolated points, so it is finite. The *complexity* $c(M)$ of a compact 3-manifold $M$ with boundary is the minimal number of vertices of a simple spine of $M$.

A *hyperbolic 3-manifold with geodesic boundary* is a complete Riemannian manifold with boundary which is locally isometric to a half-space of hyperbolic 3-space $H^3$. By Mostow’s Rigidity Theorem (see [6] for an explicit statement in the geodesic boundary case), every compact 3-manifold admits at most one hyperbolic structure with geodesic boundary, and has therefore a well-defined hyperbolic volume (if any). A useful tool for the computation of hyperbolic volumes, used in the sequel, is the Lobachevsky function $L: \mathbb{R} \to \mathbb{R}$ defined by

$$L(\omega) = -\int_0^\omega \log |2 \sin u| \, du.$$ 

To state our result we also recall that for any integer $r \geq 2$, after fixing $q_0$ in $\mathbb{C}$ such that $q_0^2$ is a primitive $r$-th root of unity, a real-valued invariant $TV_r$ for compact 3-manifolds with boundary was defined by Turaev and Viro in [14]. We consider here these invariants normalized so that $TV_r(S^3) = 1$. 


for any \( r \). Their computation involves the complex-valued quantum 6j-symbols \( \{ i \ j \ k \ l \ m \ n \} \), where \( i, j, k, l, m, n \) are half-integers, and the real-valued quantum integers
\[
[k] = \frac{q^{k} - q^{-k}}{q^{0} - q^{-1}},
\]
defined for any integer \( k \).

**Theorem 1.2.** Let \( M \in \mathcal{M}_n \). Then:

1. \( M \) is hyperbolic with geodesic boundary and its volume is given by

\[
\text{vol}(M) = n \cdot \left( 8L \left( \frac{\pi}{4} \right) - 3 \int_{0}^{\frac{\pi}{2}} \arccosh \left( \frac{\cos t}{2 \cos t - 1} \right) \, dt \right).
\]

2. \( M \) is boundary-irreducible and acylindrical.

3. Every closed embedded incompressible surface in \( M \) is parallel to the boundary.

4. \( H_1(M; \mathbb{Z}) \cong \mathbb{Z}^n \).

5. The Heegaard genus of \( M \) is equal to \( n + 1 \).

6. \( c(M) = n \).

7. The \( r \)-th Turaev-Viro invariant of \( M \) is given by

\[
\text{TV}_r(M) = \sum_{h \in \mathbb{N}, \ 0 \leq 3h \leq r-2} \left\{ \begin{array}{c} h \ n \ h \\ h \ h \ h \end{array} \right\}^n \cdot [2h + 1]^{1-n}.
\]

Before giving the Proof of Theorem 1.2, we introduce the notion of (hyperbolic) truncated tetrahedron \([7, 9, 6]\). Let \( \Delta \) be a tetrahedron and let \( \Delta^* \) be the combinatorial polyhedron obtained by removing from \( \Delta \) small open stars of the vertices. We call lateral hexagon and truncation triangle the intersection of \( \Delta^* \) respectively with a face and with the link of a vertex of \( \Delta \). The edges of the truncation triangles are called boundary edges, the other edges of \( \Delta^* \) are called internal edges. A hyperbolic truncated tetrahedron is a realization of \( \Delta^* \) as a compact polyhedron in \( \mathbb{H}^3 \), such that the truncation triangles are geodesic triangles, the lateral hexagons are geodesic hexagons, and truncation triangles and lateral hexagons lie at right angles to each other. A truncated tetrahedron is regular if all the dihedral angles along its internal edges are equal to each other. It turns out \([7, 6]\) that for every \( \theta \) with \( 0 < \theta < \pi/3 \) there exists up to isometry exactly one regular truncated tetrahedron \( \Delta^*_\theta \) of dihedral angle \( \theta \). The boundary edges of \( \Delta^*_\theta \) all have the same length \( b = b(\theta) \), and the internal edges all have the same length \( i = i(\theta) \), as shown in Fig. 3.

The volume of \( \Delta^*_\theta \) is given by (see \([13]\)):

\[
\text{vol}(\Delta^*_\theta) = 8L \left( \frac{\pi}{4} \right) - 3 \int_{0}^{\theta} \arccosh \left( \frac{\cos t}{2 \cos t - 1} \right) \, dt.
\]
As the Proof of Theorem 1.2 will now make clear, truncated tetrahedra can be used as building blocks to construct hyperbolic manifolds with geodesic boundary [7, 9, 6].

Proof of Theorem 1.2. First of all, we fix a one-edged ideal triangulation $T$ of $M$ and we denote by $e$ the single edge of $T$.

In order to give $M$ a hyperbolic structure, we identify each tetrahedron of $T$ with a copy of the regular truncated tetrahedron $\Delta_{\pi/3n}^*$. Due to the symmetries of $\Delta_{\pi/3n}^*$, every pairing between the faces of the tetrahedra of $T$ can be realized by an isometry between the corresponding lateral hexagons of the $\Delta_{\pi/3n}^*$'s. This procedure defines a hyperbolic metric on $M$, possibly with a cone singularity along $e$. But each tetrahedron is incident 6 times to $e$, so the cone angle around $e$ is $6 \cdot n \cdot \pi/3n = 2\pi$, and $M$ is actually hyperbolic without singularities. Finally, $\text{vol}(M) = n \cdot \text{vol}(\Delta_{\pi/3n}^*)$ and Point (1) is proved.

Point (2) is a general consequence of the existence of a hyperbolic structure with geodesic boundary, and Point (3) is proved using Haken’s theory of normal surfaces. If $\Sigma \subset M$ is a closed incompressible surface, then $\Sigma$ can be isotoped into normal position with respect to $T$, so the intersection of $\Sigma$ with an ideal tetrahedron $T \in T$ consists of triangles and squares. Recall now that $\Sigma$ intersects all the edges of $T$ in the same number $k$ of points. Assuming there are $q$ squares and $t_i$ triangles of type $i$, for $i = 1, \ldots, 4$, we deduce the relations $t_1 + t_2 = t_3 + t_4 = t_1 + t_3 + q = t_1 + t_4 + q = t_2 + t_3 + q = t_2 + t_4 + q = k$. This easily implies that $q = 0$ and $t_i = k/2$ (so $k$ is even), hence $\Sigma \cap T$ consists of $k/2$ parallel triangles at each vertex of $T$. Therefore $\Sigma$ consists of $k/2$ disjoint surfaces parallel to $\partial M$.

Concerning Point (4), let $P$ be the standard spine of $M$ dual to $T$. Since $M$ collapses onto $P$, we have $H_1(M; \mathbb{Z}) \cong H_1(P; \mathbb{Z})$, and we can use cellular homology to compute $H_1(P; \mathbb{Z})$. To do so, we choose a maximal tree $Y$ in the 4-valent graph $S(P)$. Then $S(P) \setminus Y$ consists of $n + 1$ edges $e_1, \ldots, e_{n+1}$.
Choose an orientation on each $e_i$ and on the single face $F$ of $P$. The face $F$ is incident three times to each $e_i$. For $i = 1, \ldots, n + 1$ let $r_i$ be the sum of a contribution $\pm 1$ over the three instances $F$ runs along $e_i$, with sign depending on whether orientations are matched or not. So $r_i \in \{-3, -1, 1, 3\}$ and $H_1(P; \mathbb{Z}) \cong \mathbb{Z}^{n+1}/(r)$ where $r = (r_1, \ldots, r_{n+1})$. We will now prove that $r_i = \pm 1$ for some $i$, which implies that $H_1(P; \mathbb{Z}) \cong \mathbb{Z}$. Let $v \in V(P)$ be an extremal vertex of $Y$, i.e., a vertex adjacent to only one edge of $Y$, whence to three of the $e_i$’s, say $e_{i_1}, e_{i_2}, e_{i_3}$ (indices could be repeated according to multiplicity of incidence). Looking at the picture of the neighbourhood of $v$ in $P$, shown in Fig. 1-(3), one now easily sees that $F$ cannot be incident three times in the same direction to each $e_{i_1}, e_{i_2}, e_{i_3}$, so we have $r_i = \pm 1$ for some $i \in \{i_1, i_2, i_3\}$.

Let us turn to Point (5). Lemma 1.1 shows that $\partial M$ has genus $n$ but, by Point (2), $M$ is not a handlebody, so its genus is at least $n + 1$. A Heegaard surface having genus $n + 1$ is simply given by the boundary of a regular neighbourhood of $\partial M \cup e$, whence the conclusion.

We now prove Point (6). The standard spine $P$ dual to $T$ has $n$ vertices, so we have $c(M) \leq n$. By Point (2), a result of Matveev [11] shows that there is a standard spine $Q$ of $M$ (not just a simple one) with precisely $c(M)$ vertices. An Euler characteristic computation gives $\chi(Q) = x - c(M)$ where $x \geq 1$ is the number of faces of $Q$, so $c(M) > 1 - \chi(Q) = 1 - \chi(M) = n$, the last equality having been proved in Lemma 1.1.

Point (7) is an easy calculation. We follow the notation of [14]: Since there is only one face $F$, a colouring of $P$ is given by assigning to $F$ a half-integer in $\{0, 1/2, 1, 3/2, \ldots, (r - 2)/2\}$. Since the same $F$ is incident three times to each edge, the colouring is admissible if it is an integer $h$ with $0 \leq 3h \leq r - 2$. Each such colouring contributes to $\text{TV}_r(M)$ with a summand given by the product of a factor $\left\{ \begin{array}{ccc} h & h & h \\ h & h & h \end{array} \right\} \cdot [2h + 1]^{-1}$ for each vertex and of a factor $[2h + 1]$ due to the single face of $P$. \hfill \Box

**Remark 1.3.** The Proof of Theorem 1.2-(1) actually shows that every one-edged ideal triangulation of $M \in \mathcal{M}_n$ is combinatorially equivalent to a decomposition of $M$ into $n$ regular truncated tetrahedra of dihedral angle $\pi/3n$.

**Remark 1.4.** It was proved in [13] that among compact hyperbolic 3-manifolds with geodesic boundary and fixed Euler characteristic $\chi < 0$, those having minimal volume are decomposed into $1 - \chi$ copies of $\Delta_{\pi/3}^{\chi/3(1-\chi)}$. Therefore $\mathcal{M}_n$ is precisely the set of hyperbolic 3-manifolds $M$ with geodesic boundary having minimal volume among compact orientable manifolds with $\chi(M) = 1 - n$. 
Remark 1.5. It follows from the Proof of Theorem 1.2-(6) that $\mathcal{M}_n$ is also the set of all hyperbolic manifolds $M$ having minimal complexity among compact orientable manifolds with $\chi(M) = 1 - n$.

Remark 1.6. Point (7) of Theorem 1.2 shows in particular that $\operatorname{TV}_r(M)$, for $M$ in $\mathcal{M}_n$, depends only on $n$ and $r$ (actually, on $q_0$), but not on $M$. This could also be proved using the fact that two standard spines with the same incidence relations between faces and vertices produce the same Turaev-Viro invariants (see [12] for more details). In fact, each manifold in $\mathcal{M}_n$ admits a spine with one face and $n$ vertices, so the incidence relations are always the same.

Remark 1.7. We believe that Theorem 1.2 extends, with minor variations, to non-orientable manifolds admitting a one-edged triangulation. However, to prove this extension, one should first generalize Matveev’s theorem [11] on standardness of minimal spines to the case of non-orientable, boundary-irreducible manifolds not containing projective planes or essential Möbius strips.

Uniqueness of the minimal spine. We are now left with two tasks: We must provide examples of manifolds in $\mathcal{M}_n$, and we must be able to distinguish manifolds in the same $\mathcal{M}_n$. We will face the former task in the next section, by constructing standard polyhedra with $n$ vertices and one face. Concerning the latter task, we have just shown that homology, Heegaard genus, Turaev-Viro invariants, and volume of manifolds in $\mathcal{M}_n$ depend on $n$ only, so we need a more powerful tool. This tool is provided by hyperbolic geometry. We recall that the cut-locus of a hyperbolic manifold $M$ with geodesic boundary is the set of all points of $M$ which admit at least two distinct distance-minimizing geodesics to $\partial M$.

Theorem 1.8. Every $M \in \mathcal{M}_n$ has a unique standard spine with $n$ vertices, homeomorphic to the cut-locus of $M$.

Proof. Let $C \subset M$ be the cut-locus of $M$, and let $P$ be a standard spine of $M$ with $n$ vertices. By Remark 1.3, $P$ is dual to a decomposition $T$ of $M$ into $n$ regular truncated tetrahedra. We claim that $C$ intersects each tetrahedron $T \in T$ as in Fig. 2-right. This implies that $C$ is homeomorphic to $P$, whence the conclusion.

To prove our claim it is sufficient to show that for every tetrahedron $T$ of $T$, every point $p$ in $T$, and every distance-minimizing geodesic $\gamma$ connecting $p$ to $\partial M$, we have that $\gamma$ is entirely contained in $T$. If this were not true, since $\gamma$ meets $\partial M$ at a right angle, a subarc $\gamma'$ of $\gamma$ would connect a truncation triangle and its opposite hexagon in some tetrahedron $T' \in T$. Then the length of $\gamma'$ would be greater than the distance between such a truncation triangle and its opposite hexagon, which in turn is greater than the distance
between \( p \) and some truncation triangle in \( T \), since \( T \) and \( T' \) are isometric to each other and regular. \(\square\)

**Remark 1.9.** An alternative proof of Theorem 1.8 could be based on the machinery developed in [6]: Remark 1.3 and the tilt-formula [16, 15, 6] easily imply that a one-edged ideal triangulation of a manifold \( M \in \mathcal{M}_n \) is combinatorially equivalent to Kojima’s canonical decomposition of \( M \), which is obtained by straightening the ideal triangulation dual to the cut-locus of \( M \) (see [9]).

We say that a standard polyhedron is *orientable* if it can be embedded in an orientable 3-manifold. Note that an orientable standard polyhedron is automatically the spine of an orientable 3-manifold, which is unique by [3].

**Corollary 1.10.** The set \( \mathcal{M}_n \) is in one-to-one correspondence with the set of orientable standard polyhedra with \( n \) vertices and one face.

**Automorphisms and chirality.** Another remarkable consequence of Theorem 1.8 is that there is an easy algorithm running in \( n^2 \) time to determine all the automorphisms of an element \( M \) of \( \mathcal{M}_n \). This algorithm will check in particular whether \( M \) is amphichiral or chiral, i.e., whether \( M \) admits an orientation-reversing automorphism or not. The algorithm is based on the following consequence of Theorem 1.8:

**Corollary 1.11.** The automorphisms of \( M \in \mathcal{M}_n \) correspond bijectively to the combinatorial automorphisms of the triangulation dual to the unique minimal spine of \( M \).

The algorithm now works as follows: First, pick an arbitrary “base” tetrahedron \( T_0 \) in the ideal triangulation \( T \) of \( M \). Second, for each \( T \) in \( T \), consider the 24 combinatorial isomorphisms \( f : T_0 \to T \) (12 of which are orientation-reversing). Third, check whether \( f \) extends to the whole of \( T \). The process of checking whether a given \( f \) extends is linear in \( n \), so the whole algorithm runs in time proportional to \( n^2 \).


The results of the previous section would of course be of little or no interest if \( \mathcal{M}_n \) (the class of manifolds having a triangulation with one edge and \( n \) tetrahedra) turned out to be empty or very small. In this section we prove that \( \#\mathcal{M}_n \) grows at least exponentially with \( n \), deducing that the number of orientable compact 3-manifolds of complexity \( n \) also grows at least exponentially (Corollary 2.6). We do this by concentrating on a special class of one-edged triangulations, and we give some hints showing that our exponential lower estimate on \( \#\mathcal{M}_n \) is actually far from being sharp.\(^1\)

\(^1\) *Added in proof.* We have actually shown in math.GT/0301114 that \( \#\mathcal{M}_n \) has growth type \( n^n \).
Oriented spines and o-graphs. In the whole of this section we consider oriented (rather than just orientable) manifolds (but recall that chirality of the elements of $\mathcal{M}_n$ can be checked very easily). We remind [1] that if $P$ is a standard spine of an oriented $M$ then $P$ also carries an orientation, defined as a screw-orientation along the edges of $S(P)$ with a natural compatibility at vertices (see [2, Fig. 2]). Conversely, if $P$ is an oriented standard polyhedron, then $P$ is orientable, and the manifold it defines is oriented. In addition, $P$ can be described by two additional structures on the 4-valent graph $S(P)$:

- An embedding in the plane of the neighbourhood of each vertex, with two opposite strands marked as being over the other two, as in knot projections.
- A colour in $\mathbb{Z}/3$ attached to each edge.

A 4-valent graph with these additional structures is called an o-graph. It was shown in [1] that any o-graph defines an oriented standard polyhedron, whence an oriented manifold, and that two o-graphs defining the same oriented polyhedron are related by certain "C-moves." The effect of a C-move is to change the planar structure at a vertex and the $\mathbb{Z}/3$-colouring of the edges incident to this vertex.

O-graphs based on the open chain. Let $G_n$ be the graph with vertices $v_1, \ldots, v_n$, a closed edge at $v_1$ and one at $v_n$, and two edges joining $v_i$ to $v_{i+1}$ for $i = 1, \ldots, n - 1$. We characterize in this paragraph the oriented standard polyhedra $P$ such that $S(P) = G_n$ and $P$ has a single face. We begin with the following fact that one can readily establish using the C-moves mentioned above:

**Lemma 2.1.** Any oriented standard polyhedron $P$ such that $S(P) = G_n$ can be represented by an o-graph as shown in Fig. 4.

![Figure 4. O-graph of a generic polyhedron based on $G_n$.](image)

The description of which o-graphs as in Fig. 4 have a single face will use the language of finite state automata (see for instance [4]). We recall that a finite state automaton over a finite set $\mathcal{A}$ (the alphabet) consists of a finite set $\mathcal{S}$ (the states), a function $\mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$ (the transition function), an element $s_0$ of $\mathcal{S}$ (the start state), and a subset $\mathcal{S}'$ of $\mathcal{S}$ (the set of accept states). A word (a finite string of letters from the alphabet) is accepted
by the automaton if, starting from $s_0$, reading the word from left to right, and using the transition function, the automaton ends in a state of $S'$. An automaton can be encoded by a picture, where the states are represented by $S$-labeled boxes (with double margin for accept states), the transition function is given by box-to-box $A$-labeled arrows, and a mark indicates the start state.

**Theorem 2.2.** The oriented standard polyhedron defined by the $o$-graph of Fig. 4 has a single face if and only if both $\alpha$ and $\delta$ are different from 2 and the word

$$(\beta_1, \gamma_1)(\beta_2, \gamma_2) \ldots (\beta_{n-1}, \gamma_{n-1})$$

is accepted by the automaton described in Fig. 5, where

$A_0 = \{(2, 2)\}, \quad A_1 = \{(0, 0), (1, 1)\}, \quad A_2 = \{(1, 0), (0, 1)\},$

$A_3 = \{(0, 2), (1, 2), (2, 0), (2, 1)\}, \quad A = A_0 \cup A_1 \cup A_2 \cup A_3.$

**Proof.** We confine ourselves to a general explanation, omitting the many combinatorial details. We analyze the graph left to right, starting from $\alpha$.

![Figure 5. A finite state automaton with alphabet $A = A_0 \cup A_1 \cup A_2 \cup A_3$.](image)

Figure 5 shows that for $\alpha = 2$ there are at least two faces, so $\alpha \in \{0, 1\}$. Now we examine the pair of colours $(\beta_1, \gamma_1)$ starting with the string $XYYYX$.  

![Figure 6. Portions of o-graph.](image)
which describes the way the faces already constructed are matched to the left of the point we are considering. (Note that both \( \alpha = 0 \) and \( \alpha = 1 \) give \( XYYX \).) Depending on \((\beta_1, \gamma_1)\) we will have either the creation of a face which does not fill the polyhedron, or a new pattern that describes the matching of faces. More generally, as we proceed, we will have a string \( xyzw \) of two symbols each repeated twice, and we will have to analyze the effect of a pair of colours \((i, j)\), which either creates a closed face or produces a new pattern \( x'y'z'w' \) (see Fig. 6 again). The detailed analysis of all the possibilities leads precisely to the transitions shown in Fig. 5, where “FAIL” means that a closed face is created. To conclude we note that the final edge with colour \( \delta \in \{0, 1\} \) gives a single global face when the input pattern is \( XYYY \) or \( XYXY \), and not otherwise.

□

Growth of \( \#M_n \). Theorem 1.8 shows that to compute \( \#M_n \) it is sufficient to count the combinatorially distinct standard polyhedra with \( n \) vertices and one face. Restricting to the oriented ones with the open chain \( G_n \) as singular graph we must then discuss which \( \alpha \)-graphs as in Fig. 4 define the same polyhedron. A move that of course does not change the polyhedron associated to the \( \alpha \)-graph is the 180° degree rotation. Using the C-moves of [1] one can see that another such move consists in interchanging each \( \beta_k \) with the corresponding \( \gamma_k \). In addition, these two moves are sufficient to generate all graphs giving the same polyhedron. Therefore we have the following:

Lemma 2.3. An oriented standard polyhedron is defined by at most four different \( \alpha \)-graphs as in Fig. 4.

Proposition 2.4. There are at least \( 4 \cdot 12^{(2n-5)/3} \) distinct oriented standard polyhedra with one face and the open chain with \( n \) vertices as singular set.

Proof. We must count the possible choices for the \( \beta_k \)'s and \( \gamma_k \)'s, i.e., the words of length \( n - 1 \) accepted by the finite state automaton of Fig. 5, multiply by 4 (the choices for \( \alpha \) and \( \delta \)), and divide by at most 4 according to the previous lemma. So it is sufficient to prove that there are at least \( 4 \cdot 12^{(2n-5)/3} \) words of length \( n - 1 \) accepted by the automaton.

The idea is just to perform the loop \( XYYX \to XYYY \to XYXY \to XYXY \) in all possible ways, inserting a single loop \( XYYX \to XYYY \) when \( n - 1 \) is a multiple of 3 (for in this case we would end up in the non-accept start state). Since 6 letters lead from \( XYYX \) to \( XYYY \), 4 lead from \( XYYY \) to \( XYXY \), and again 6 from \( XYXY \) to \( XYYX \), it is clear that approximately \( 6^2(n-1)/3 \cdot 4(n-1)/3 = 12^{2(n-1)/3} \) words can be constructed with this method. The exact computation carried out depending on the congruence class of \( n - 1 \) modulo 3 leads to the desired estimate. □
The next easy remark shows that the qualitative type of growth just established is actually the maximal one could expect. After the remark we also give an obvious consequence of the previous proposition.

**Remark 2.5.** Given a 4-valent graph $G$ with $n$ vertices, there exist at most $2^n \cdot 3^{2n} = 18^n$ distinct oriented standard polyhedra $P$ such that $S(P) = G$. If $G = G_n$, using Lemma 2.1, one can see that there exist at most $3^{2n} = 9^n$ of them.

**Corollary 2.6.** There exist $c > 0$ and $b > 1$ such that $\# M_n \geq c \cdot b^n$. In particular, the number of distinct orientable, boundary-irreducible, and acylindrical manifolds of complexity $n$ is at least $c \cdot b^n$.

We remind the reader that the manifolds referred to in the previous corollary are precisely those known to have standard minimal spines. Now we have:

**Remark 2.7.** The number of distinct oriented standard polyhedra with $n$ vertices is bounded from above by $18^n \cdot g(n)$, where $g(n) \leq (4n - 1)!!$ is the number of distinct four-valent graphs with $n$ vertices.

**Further comments on estimates.** The lower bound on the number of elements of $\mathcal{M}_n$ based on the open chain graph $G_n$ provided by Proposition 2.4 is very far from being sharp. For instance, if we consider in Fig. 5 the paths consisting of some $XYXY \rightarrow XXYY \rightarrow XYYX \rightarrow XYXY$ cycles intermingled with some $XXYY \rightarrow XYYX$ loops, we deduce that the number of distinct spines is at least

$$
\sum_{k=0}^{n-2} \begin{cases} 
0 & \text{if } n - k - 1 = 3h, \\
2^k \cdot 6^{2h+1} \cdot 4^h \cdot \binom{h+k}{h} & \text{if } n - k - 1 = 3h + 1, \\
2^k \cdot 6^{2h+1} \cdot 4^{h+1} \cdot \binom{h+k}{h} & \text{if } n - k - 1 = 3h + 2.
\end{cases}
$$

Concentrating on the term of the sum corresponding to $k = \lfloor n/7 \rfloor$ and using Stirling’s formula one can for instance deduce from this estimate that there exists $c > 0$ such that, for all $\varepsilon > 0$, the number of oriented elements of $\mathcal{M}_n$ based on $G_n$ is at least $c \cdot (6 - \varepsilon)^n$ for $n \gg 0$. Note that $12^{2/3} \approx 5.2418$.

**Remark 2.8.** With the aid of a computer, in [10] we have listed and classified the approximately 2,000 closed, irreducible and orientable manifolds of complexity up to 9. Corollary 2.6 suggest that a similar listing for orientable, compact, boundary-irreducible, acylindrical manifolds may be hopeless. Our lower bound on their number, even if not sharp, already implies that there are at least 115,000 such manifolds of complexity up to 9.
Another example: The closed chain. We have concentrated in this section on the open chain $G_n$, because this graph is already sufficient to show that $\#\mathcal{M}_n$ grows exponentially. But we believe that a systematic investigation of the 4-valent graphs supporting a polyhedron with a single face would be quite interesting. Our guess is actually that most graphs indeed support many different such polyhedra. As another example we only mention the closed chain, shown in Fig. 7.

![Figure 7. An o-graph based on the closed chain.](image)

The combinatorial analysis is in this case harder than that carried out for Theorem 2.2, but we state at least the following fact, which already implies that again in this case the number of relevant polyhedra grows exponentially with $n$:

**Proposition 2.9.** Fix $(\alpha_1, \beta_1) \in \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ in the o-graph of Fig. 7. Let $(\alpha_2, \beta_2) \ldots (\alpha_n, \beta_n)$ be a random word $w$ in the letters $\{(0, 2), (1, 2), (2, 0), (2, 1)\}$. Then for $n \gg 0$ there is a probability $1/2$ that $w$ defines a polyhedron with a single face.

Chirality. We get back here to the chirality issue discussed at the end of Section 1, showing how o-graphs can be used in connection with it. We first recall [1] that if an o-graph $\Gamma$ represents an oriented manifold $M$ then the manifold $-M$ obtained by reversing the orientation of $M$ is represented by the o-graph $-\Gamma$ obtained from $\Gamma$ by switching overstrands and understrands at vertices, and changing each edge-colour to its opposite in $\mathbb{Z}/3$. Now assume $M$ belongs to $\mathcal{M}_n$ and $\Gamma$ represents the unique minimal spine of $M$. Then $M$ is amphichiral if and only if $-\Gamma$ defines the same oriented polyhedron as $\Gamma$, which can be checked in $n^2$ time using an algorithm based
on C-moves. Turning to the $M$’s based on $G_n$ and using the C-moves again, one can now prove the following:

**Proposition 2.10.** If $\Gamma$ is the o-graph of Fig. 4 then $-\Gamma$ is a similar o-graph with colours:

$$
\alpha' = 1 - \alpha, \quad \beta'_{2k+1} = 1 - \gamma_{2k+1}, \quad \beta'_{2k} = 1 - \beta_{2k},
$$

$$
\delta' = 1 - \delta, \quad \gamma'_{2k+1} = 1 - \beta_{2k+1}, \quad \gamma'_{2k} = 1 - \gamma_{2k}.
$$

This proposition and Lemma 2.3 show that chirality for the elements of $\mathcal{M}_n$ obtained from the open chain $G_n$ can be tested in linear time: We only need to check whether one of four given $2n$-tuples of elements of $\mathbb{Z}/3$ coincides with another such given $2n$-tuple.

**References**


Received July 3, 2002 and revised September 14, 2002.

Scuola Normale Superiore
Piazza dei Cavalieri 7
56127 Pisa
Italy
E-mail address: frigerio@sns.it

Dipartimento di Matematica
Università di Pisa
Via F. Buonarroti 2
56127 Pisa
Italy
E-mail address: martelli@mail.dm.unipi.it

Dipartimento di Matematica Applicata
Università di Pisa
Via Bonanno Pisano 25B
56126 Pisa
Italy
E-mail address: petronio@dm.unipi.it
HIGHER EXCHANGE RELATIONS

Shamindra Kumar Ghosh

In this paper, we extend Landau’s notion of ‘exchange relations’ so as to make sense for arbitrary planar algebras, which need not necessarily be generated by its ‘2-boxes’. We show, as in Landau’s case, that these ‘higher exchange relation planar algebras’ are necessarily ‘finite dimensional’, and that examples of such planar algebras are given by all (even possibly reducible) depth two subfactors, as well as planar algebras associated to subfactors with principal graphs $E_6$ and $E_8$.

1. Introduction

In [Jon1], V.F.R. Jones introduced the notion of index for Type II$_1$ subfactors and also examined the canonical tower,

\[ N \subset M \subset M_1 \subset M_2 \subset \cdots \]

of basic construction of $N \subset M$. For a finite index inclusion of Type II$_1$ factors, $N \subset M$, the grid of finite dimensional algebras of relative commutants,

\[
\begin{align*}
N' \cap N, & \quad N' \cap M, & \quad N' \cap M_1, & \quad N' \cap M_2, & \quad \cdots \\
\cup & \quad \cup & \quad \cup & \quad \cup & \quad \\
M' \cap M, & \quad M' \cap M_1, & \quad M' \cap M_2, & \quad \cdots
\end{align*}
\]

known as the standard invariant, became an important invariant for $N \subset M$ (see [GHJ], [JS], [Pop1], [Pop2]).

Popa ([Pop2]) studied the question of which families \( \{A_{ij} : -1 \leq i \leq j < \infty \} \) of finite-dimensional $C^*$-algebras could arise as the tower of relative commutants of an extremal finite-index subfactor — i.e., when does there exist such a subfactor $M_{-1} \subset M_0$ such that $A_{ij} = M'_i \cap M_j$; and he obtained a beautiful algebraic axiomatisation of such families, which he called $\lambda$-lattices. Subsequently, Jones used this characterisation of $\lambda$-lattices to obtain an algebraic and geometric reformulation of the standard invariant, which he called Planar Algebras (see [Jon2]).

In [Jon2], a construction of planar algebras by generators and relations is described. In [Lan], a condition called ‘the exchange relation’ is obtained, which guarantees that a planar algebra ‘generated by finitely many of its 2-boxes’ satisfies the crucial requirement of finite dimensionality. In this paper we extend and generalise the ‘exchange relation’ condition in [Lan]; we relax
the ‘2-box’ requirement of the generating set in [Lan] and define a ‘higher exchange relation’. Our definition does not demand the planar algebra to be irreducible (as in [Lan]); but puts the restriction that the ATL₁-submodule (in the sense of [Jon3]) generated by the generators of the planar algebra is closed under multiplication. We show that such ‘higher exchange relation planar algebras’ are indeed finite dimensional (see Theorem 6). The ideas of the proof of Lemma 5 and Theorem 6 have a similar flavour to that of Theorem 1 in [Lan], namely, the ‘internal face’ consideration and the use of Euler characteristic formula. However, the computations are more involved due to the fact that the planar algebras are generated by ‘k-boxes’ instead of ‘2-boxes’. Essential to these computations are the notions that we have termed ‘i-capped k-boxes’ and ‘proper internal face’ (which latter notion was introduced instead of the ‘internal face’ of [Lan], where only the irreducible case is treated). Proposition 8 shows that any depth two subfactor corresponds to a higher exchange relation planar algebra; this generalizes Theorem 4 in [Lan], where only irreducible depth two subfactors were considered. (It should be noted that an exchange relation planar algebra in the sense of [Lan] is also a higher exchange relation planar algebra.) We conclude by showing that the planar algebras associated to the $E_6$ and $E_8$ subfactors are higher exchange relation planar algebras.

2. Planar algebra.

We recall the notion of planar algebra in the form of generators and relations as described in [Jon2]. We first give some definitions.

For $k \geq 0$, by a standard $k$-box, we will mean a rectangle — or sometimes a circle — with $2k$ points which are numbered in a clockwise fashion.

![Figure 1](image)

We also need a labelling set $L = \bigsqcup_{k=0}^{\infty} L_k$ — where $\bigsqcup$ denotes disjoint union. A standard $k$-box will be said to be $L$-labelled if it has been assigned an element of $L_k$. 
Definition 1. An $L$-labelled $k$-tangle, for $k \geq 1$ consists of:

(a) A standard $k$-box which forms the external boundary,
(b) finitely many (may be zero) $L$-labelled standard $l$-boxes inside the external standard $k$-box, for $l \geq 1$,
(c) oriented planar strings, some of which may constitute loops, the others connecting two boxes in such a way that each of the marked points on the boxes is attached to exactly one string which ends at some other marked point, and the orientation of the strings satisfies the following rules:

(i) A string attached to an odd point of the external box or an even point of an internal box is oriented towards the point;
(ii) a string attached to an even point of the external box or an odd point of an internal box is oriented away from the point.

(We are given a checkerboard shading of the connected components of the complement of the strings and internal boxes inside a $k$-tangle, whereby, as one moves along the orientation of any string, the region to the right (resp., left) is shaded (resp., unshaded).)

We consider tangles up to planar isotopy. So, an $L$-labelled $k$-tangle is an equivalence class of pictures defined above under planar isotopy of the strings. Figure 2 shows an example of $L$-labelled 5-tangle with two internal boxes, and $l \in L_2$, $m \in L_3$. After shading, the tangle looks like Figure 3.

In Figure 3, the *'s, present in the white region near the boundary of the external disc and the two internal discs, indicate the white region adjacent to the first point in each disc. Similarly, in any tangle, instead of numbering each point of the external or internal boxes, we can just put a * near the boundary of the box in the white region adjacent to the first point.
On the other hand, let us define an $L$-labelled $0$-tangle by Conditions (a), (b), (c) of the above definition for $k = 0$, as well as the requirement that the connected component adjacent to the external boundary is unshaded.

Let $T_k(L)$ denote the set of all $L$-labelled $k$-tangles and let $P_k(L)$ be the vector space with $T_k(L)$ as a basis. Then $P_k(L)$ has a natural algebra structure, where basis elements — i.e., tangles — are multiplied by juxtaposition (as with braids). From the braid analogy, it is no surprise that (a) the identity element of $P_k(L)$ is given by the identity tangle where there are $k$ strings going ‘straight down’, and that (b) $P_k(L)$ may be identified with the subalgebra of $P_{k+1}(L)$ in such a way that an element of $T_k(L)$ is thought of as an element of $T_{k+1}(L)$ whose ‘right-most strand comes straight down’.

$\bigcup_{k \geq 0} P_k(L)$ is called the Universal Planar Algebra on $L$, and is denoted by $P(L)$. 

$1 \in P_0(L) \subset P_1(L) \subset P_2(L) \subset \cdots \bigcup_{k \geq 0} P_k(L)$. 

So we have a chain of unital algebras, namely,
Next, we recall the notion of annular action. By an $L$-labelled $j$-$k$-annular tangle, we shall mean a $k$-tangle with a distinguished unlabelled internal $j$-box, all other internal boxes being $L$-labelled. Let $A_{j,k}(L)$ denote the set of all $L$-labelled $j$-$k$-annular tangles. Each $A \in A_{j,k}(L)$ defines a natural linear map $A : \mathcal{P}_j(L) \to \mathcal{P}_k(L)$ thus: If $T \in \mathcal{T}_j(L)$, then $A(T)$ is defined as the $k$-tangle obtained by gluing the external boundary of $T$ with the boundary of the distinguished $j$-box of $A$ in such a way that the marked points of both these boundaries ‘match’ under the gluing; the action of $A$ is then extended linearly over $\mathcal{P}_j(L)$.

It should be clear that this defines an ‘action’ in the sense that if also $B \in A_{k,m}(L)$, then, the result of ‘gluing’ $A$ into the distinguished internal $k$-box of $B$ — taking care to glue pairs of identically marked points — yields an element $B \circ A \in A_{j,m}(L)$; and

$$(B \circ A)(T) = B(A(T)).$$

We are finally ready to define $P\langle L, R \rangle$, the general planar algebra with generators $L$ and relations $R$.

First fix a subset $\mathcal{R} \subset \mathcal{P}(L)$ and define $\mathcal{I}_k(\mathcal{R}) = \text{span}\{A(X) : A \in A_{j,k}(L), X \in \mathcal{R} \cap \mathcal{P}_j(L), j \geq 0\}$. It turns out (see [Jon2]) that $\mathcal{I}_k(\mathcal{R})$ is an ideal in $\mathcal{P}_k(L)$, and that $\mathcal{I}_{k+1}(\mathcal{R}) \cap \mathcal{P}_k(L) = \mathcal{I}_k(\mathcal{R})$. Therefore, $\mathcal{I}(\mathcal{R}) = \bigcup_{k \geq 0} \mathcal{I}_k(\mathcal{R})$ is an ideal in $\mathcal{P}(L)$.

Consider the quotient map $\Phi_{\mathcal{R}} : \mathcal{P}(L) \to \frac{\mathcal{P}(L)}{\mathcal{I}(\mathcal{R})} = P\langle L, R \rangle$. Define $P_k\langle L, R \rangle = \frac{\mathcal{P}_k(L)}{\mathcal{I}_k(\mathcal{R})} = \Phi_{\mathcal{R}}(\mathcal{P}_k(L))$. Thus we have another chain of unital algebras:

$$P_0\langle L, R \rangle \subset P_1\langle L, R \rangle \subset P_2\langle L, R \rangle \subset \cdots \bigcup_{k \geq 0} P_k\langle L, R \rangle = P\langle L, R \rangle.$$  

We say $P\langle L, R \rangle$ is a planar algebra\(^1\) if it satisfies the following two conditions:

(i) $$\dim(P_0\langle L, R \rangle) = 1 = \dim(P_{1,1}\langle L, R \rangle),$$

(ii) both $$\Phi_{\mathcal{R}}(\raisebox{-.5cm}{\includegraphics[width=.5cm]{figure1.png}}) \text{ and } \Phi_{\mathcal{R}}(\raisebox{-.5cm}{\includegraphics[width=.5cm]{figure2.png}})$$

are nonzero.

\(^1\)It must be remarked that in order to make contact with subfactors — like Jones — it is necessary to also require that what we have called a planar algebra should satisfy further conditions so as to become ‘finite-dimensional, spherical and $C^*$-planar algebras’.
In Condition (i) of the above definition, we have, following [Jon2], used the notation $P_{1,1}(L, \mathcal{R}) = \Phi_{\mathcal{R}}(P_{1,1}(L))$ and

$$P_{1,1}(L) = \text{span}\left\{ T \in P_1(L) \mid \text{the two points of the boundary of } T \text{ are connected by a string} \right\}.$$ 

The map $\Phi_{\mathcal{R}}$ is said to present the planar algebra $P(L, \mathcal{R})$ on the labelling set $L$ with relations $\mathcal{R}$.

3. Higher exchange relation planar algebra.

In this section, we will define what we mean by a ‘higher exchange relation planar algebra’ and we will show that any finitely generated planar algebra which satisfies this condition is ‘locally finite-dimensional’. (This will generalise the ‘exchange relation’ condition of [Lan] which, however, makes sense only for planar algebras generated by finitely many 2-boxes.)

Begin by noting that an ‘annular $j$-$k$-tangle on the empty set’ (i.e., an element of $A_{j,k}(\phi)$) is nothing but a $k$-tangle with no internal box other than the distinguished $j$-box. Now if $1 \leq i \leq k$, define an $i$-capped $k$-box to be a $(k - i)$-box obtained by applying any annular tangle in $A_{k-i,k}(\phi)$ on a $k$-box; this has the effect of putting $i$ caps on the $k$-box. Note that there will be exactly $2(k - i)$ many free strings coming out of an $i$-capped $k$-box. We shall write $l^{(i)}$ to denote an $i$-capped $k$-box whose internal $k$-box is labelled by $l$.

Figure 5 gives two examples of 2-capped 3-boxes.

![Figure 5](image)

Remark 2. Before proceeding to the next (and crucial, for us) definition of a ‘higher exchange relation’, which might be a little forbidding at first glance, it will help if we digress briefly with a comparison of Landau’s original definition of an ‘exchange relation’ with our Definition 3.

(a) He works with the case where $P(L, \mathcal{R})$ is generated by its 2-boxes — i.e., $L = L_2$ and $L_k = \phi$ for $k \neq 2$ — where $\phi$ denote the empty set, and where $L$ is a finite set. We work with the general finitely presented case. Without loss of generality, we may assume that there exists a $k$ such that $L_{k'} = \phi$ if $k' \neq k$ since if $l' \in L_j$ and $j \leq k$, we may take the picture in Figure 6 as a new label, $l \in L_k$. 

![Figure 6](image)
(b) He needs (and implicitly uses, although he does not explicitly include this in his definition) the fact that in his planar algebra, ‘contractible loops can be replaced by scalars’. We explicitly assume this as Condition (o) of Definition 3.

(c) He works with the case of ‘irreducible subfactors’, which amounts, in our case, to working with the case where

$$P_0(L, R) = P_1(L, R) = \mathbb{C}.$$  

Our Conditions (i) and (ii) of Definition 3 are appropriate generalisations of these two requirements. When $k = 2$, our conditions are equivalent to demanding that $P_0(L, R) = \mathbb{C}$ and that $P_1(L, R)$ is finite-dimensional.

(d) The really important condition in Landau’s definition is the last condition, the one which motivated his terminology (on the basis of the ‘group example’). This condition stems from the fact that there are essentially three different kinds of 3-tangles which have two internal 2-boxes connected by one string; and the condition requires that any labelled tangle of one of these kinds is expressible as linear combinations of labelled tangles of the other 2-kinds. The needed generalisation of this condition requires the use of the ‘rotations’ which play an all-important role in Jones’ theory of planar algebras. (Recall that for each $j \geq 1$, the rotation is the annular $j$-$j$ tangle $\rho_j \in \mathcal{A}_{j,j}(\phi)$ where the point marked 1 on the internal disc is connected to the point marked 3 on the external disc.) According to [Jon2], the map $T \mapsto \rho_j(T)$ defines a bijective self-map of $P_j(L, R)$ which, in the subfactor case, preserves the ‘inner-product induced by the trace’.

**Definition 3.** $P(L, R)$ is said to be a **higher exchange relation planar algebra** if $R$ contains the following elements:

(o) If $\gamma$ is a contractible loop, then

$$\left( \gamma - \delta 1 \right) \in R$$  

for some scalar $\delta$. 

*Figure 6.*
(i) For any \( k \)-capped labelled \( k \)-box, \( l^{(k)} \) (with \( l \in L = L_k \)), there exists a constant \( C(l^{(k)}) \) such that

\[
l^{(k)} - C(l^{(k)})1 \in \mathcal{R}.
\]

(ii) For any three \( (k-1) \)-capped labelled \( k \)-boxes, \( l^{(k-1)} \), \( m^{(k-1)} \) and \( n^{(k-1)} \), there exists constants \( C_{l^{(k-1)},m^{(k-1)},n^{(k-1)}} \) such that

\[
\left( l^{(k-1)},m^{(k-1)} - \sum_{n^{(k-1)}} C_{l^{(k-1)},m^{(k-1)},n^{(k-1)}} n^{(k-1)} \right) \in \mathcal{R}.
\]

(This relation says that \( \text{span}\{\Phi_{\mathcal{R}}(l^{(k-1)}) : l \in L\} \) is closed under multiplication in \( P(L, \mathcal{R}) \).)

(iii) Let \( A \) denote any \( (2k - 1) \)-tangle with exactly two internal \( k \)-boxes, both unlabelled, which are connected by exactly one string, in which all other strings have exactly one end point on the external disc, but whose three \( *'s \) (one for the external and two for the internal boxes) can be in arbitrary places (modulo the shading requirement).

For each \( A \) above and any \( l, m \in L = L_k \) we demand that \( A(l, m) \) can be written as the sum of two terms: The first term is a suitable linear combination of \( L \)-labelled \( (2k - 1) \)-tangles obtained by applying nontrivial rotations on the external boundary of \( A \) and arbitrary rotations on the two internal boxes which can assume any labels from \( L \), the second term is a suitable linear combination of \( (2k - 1) \) tangles with at most one labelled \( k \)-box. More precisely, if we denote by \( \Lambda \), the set of all \( (2k - 1) \) tangles which have at most one \( L \)-labelled internal...
box, then we require the existence of constants $C_{x,y,i,p,q}$ and $C^{(A,l,m)}_T$ such that

$$A(l,m) = \left( \sum_{x,y \in L, 1 \leq p,q \leq k, 1 \leq i \leq (2k-2)} C^{(A,l,m)}_{(x,y,i,p,q)} \rho^{i}_{2k-1}(A(\rho^p_k(x), \rho^q_k(y))) + \sum_{T \in \Lambda} C^{(A,l,m)}_T T \right) \in \mathcal{R}.$$ 

**Remark 4.**

(a) Note that in Condition (iii), the symbol $i$ runs from 1 to $(2k-2)$ which means that we consider only nontrivial rotations of the external boundary, whereas we also allow the trivial rotation (i.e., the identity map) of the internal discs.

(b) In Condition (iii), instead of making $\ast$ for the external boundary of $A$ arbitrary, we can just demand the fact for a fixed location of the $\ast$ of the external boundary. This is because the relation for other positions of $\ast$ can be deduced by applying nontrivial rotations to the external boundary of $A$.

The interior of each tangle is partitioned by the strings and the internal boxes; the connected components will be called faces. Each face which does not touch the external boundary of the tangle will be called an internal face and the size of an internal face is defined as the number of internal boxes touched by the face. An internal face will be said to be a proper internal face, if its size is strictly greater than 1.

![Diagram of Proper Internal Face](image)

**Figure 8.**

**Lemma 5.** If $P = P(L, \mathcal{R})$ is a higher exchange relation planar algebra, and if $T \in \mathcal{T}_n(L)$, $n \geq 0$, then $\Phi_{\mathcal{R}}(T)$ can be written as a linear combination of terms of the form $\Phi_{\mathcal{R}}(T')$ where $T' \in \mathcal{T}_n(L)$ and does not contain any proper internal face.
Proof. Suppose not. Let
\[ C_1 = \left\{ T \in \mathcal{T}_n(L) : \begin{array}{l}
T \text{ does not satisfy the lemma}, \\
T \text{ has no contractible loops,} \\
\text{and has no } k\text{-capped } k\text{ boxes}
\end{array} \right\}. \]

Then \( C_1 \neq \emptyset \), by assumption and by relations (o), (i) of Definition 3. Choose \( T_1 \in C_1 \) such that number — call it \( n_1 \) — of internal boxes in \( T_1 \) is minimal among all elements \( T \in C_1 \).

Let \( C_2 = \{ T \in C_1 : \text{number of internal boxes in } T = n_1 \} \). \( C_2 \neq \emptyset \) since it contains \( T_1 \). Choose \( T_2 \in C_2 \) such that the size of its smallest proper internal face — call it \( n_2 \) — is minimal among all elements \( T \in C_2 \). (This makes sense because each element of \( C_1 \) has at least one proper internal face.) By definition of proper internal face, \( n_2 \geq 2 \). Observe that inside \( T_2 \), we can find a picture of Type \( A(l, l') \) (as in Definition 3(iii)), so that the string connecting the two discs is on the boundary of a proper internal face of \( T_2 \) of size \( n_2 \). (See Figure 9.)

![Figure 9](image-url)

Using relation (iii), \( A(l, l') \) can be replaced by a linear combination of certain pictures and thereby \( \Phi_R(T_2) \) can be written as a linear combination of \( \Phi_R(T') \)'s where \( T' \in \mathcal{P}_n(L) \). The crucial point is that while we write \( \Phi_R(T_2) \) as a linear combination of \( \Phi_R(T') \)'s via expressing \( A(l, l') \) as a linear combination of certain pictures, each \( T' \) is seen to necessarily satisfy one of the following two conditions:

(i) The number of internal boxes of \( T' \) is strictly less than \( n_1 \); or
(ii) \( T' \) has exactly \( n_1 \) internal boxes and the minimal size of its proper internal faces is strictly less than \( n_2 \).

Finally observe that the assertion of the lemma is indeed satisfied by \( T' \)'s of either of the above types, by the definitions of \( n_1 \) and \( n_2 \). Thus, the assertion of the lemma is satisfied by all \( T' \)'s which appear in the linear
combination yielding \( \Phi_R(T_2) \). Hence, \( T_2 \) must also satisfy the claim. This is a contradiction and completes the proof of Lemma 5. \( \square \)

**Theorem 6.** If \( P (= P(L, R)) \) is a higher exchange relation planar algebra, then \( P \) is a locally finite-dimensional planar algebra — i.e., \( P_k \) is finite dimensional for all \( k \).

**Proof.** We first prove that \( P_0 \) has dimension 1. It is enough, by Lemma 5 to prove that \( \Phi_R(T) \in \mathbb{C} \cdot 1 \) for \( T \in T_0(L) \) such that \( T \) does not have any proper internal face. Consider such a \( T \). Note that \( T \) cannot have two internal boxes connected by a string between them. Otherwise, there would always exist an internal face of size \( \geq 2 \), because the boundary of \( T \) is not connected to any string. Thus all internal \((k-)\)boxes in \( T \) will be \( k \)-capped.

So, \( T \) will consist of \( k \)-capped internal boxes and contractible loops; then, using the relations, (\( o \)) and (\( i \)), we find that \( \Phi_R(T) \in \mathbb{C} \cdot 1 \).

Similarly, we can prove that \( P_{1,1} = \mathbb{C} \cdot 1 \) Thus, \( P \) is a planar algebra. It remains to show that \( \dim(P_n) < \infty \).

Suppose \( T \in P_n(L) \) and \( T \) does not contain any proper internal face, contractible loop or \( k \)-capped box. Let \( b^T \) be the number of internal boxes in \( T \) and \( b^T_i \) be the number of \( i \)-capped boxes in \( T \). Clearly, \( b^T = \sum_{i=0}^{k-1} b^T_i \).

We construct a graph (embedded in the 2-sphere) from \( T \) thus: Embed the tangle \( T \) in the 2-sphere (thought of as the one-point compactification of the plane); join the point at \( \infty \) on the sphere to each of the boundary points of the tangle \( T \); and shrink each of the \( i \)-capped boxes along with their \( i \) caps to points.

Let \( V, E \) and \( F \) be the number of vertices, edges and faces of this graph. Then,

\[
V = 2n + b_T + 1
\]

\[
E = \frac{2(2n) + 2n + \sum_{i=0}^{k-1} 2(k - i) b^T_i}{2}
\]

\[
= 3n + \sum_{i=1}^{k} i b^T_{k-i}
\]

\[
F \leq 2n \text{ since } T \text{ does not have any proper internal face.}
\]

Since the Euler characteristic of \( S^2 \) is 2, we find that

\[
V - E + F = 2
\]

\[
\Rightarrow (E - V) = (F - 2)
\]

\[
\Rightarrow (E - V) \leq (2n - 2)
\]

\[
\Rightarrow (n - 1) + \sum_{i=1}^{k} (i - 1) b^T_{k-i} \leq (2n - 2)
\]

\[
\Rightarrow \sum_{i=2}^{k} (i - 1) b^T_{k-i} \leq (n - 1)
\]
\[ \sum_{i=2}^{k} b_{k-i}^T \leq (n - 1) \]

\[ (b^T - b_{k-1}^T) = \left( \sum_{i=1}^{k} b_{k-i}^T \right) - b_{k-1}^T \leq (n - 1). \]

Let \( D_1 = \left\{ T \in \mathcal{T}_n(L) : \right. \)

- \( T \) has no contractible loops,
- no \( k \)-capped box,
- no \( (k-1) \)-capped box
- and \( (b^T - b_{k-1}^T) \leq (n - 1) \).

Clearly, \( |D_1| < \infty \), since the labelling set is finite and since there exist only finitely many \( n \)-tangles with a fixed number of boxes. For \( T \in D_1 \), let us write \( T^a \) for the result obtained by inserting exactly one \( (k-1) \)-capped labelled \( k \)-box on each of some (maybe all) of the strings of \( T \). (There may be several \( T^a \)'s for each \( T \in D_1 \), but the number of such \( T^a \) is finite, because there are only a finite number of strings in \( T \) and only finitely many \( (k-1) \)-capped labelled \( k \)-boxes.) Hence, the set \( D_2 = \{ T^a : T \in D_1 \} \) is finite.

We finish the proof of the theorem by showing that \( D_2 \) linearly spans \( P_n \). Now \( P_n \) is generated by the set of \( \Phi_R(T)'s \), where \( T \) ranges over those \( n \)-tangles which have no proper internal faces, contractible loops, or \( k \)-capped boxes. Take such a \( T \). Replacing the \( (k-1) \)-capped boxes in \( T \) by ‘straight strings’, we get an element of \( D_1 \). Thus \( T \) is basically an element of \( D_1 \), on some of whose strings, one or more \( (k-1) \)-capped boxes have been inserted. By virtue of relation (ii), we can express \( \Phi_R(T) \) as a linear combination of \( \Phi_R(T')'s \), where \( T' \in D_2 \). Thus \( \dim(P_n) < \infty \). \( \square \)

**Remark 7.** If we consider a higher exchange relation planar algebra with the labelling set \( L = L_2 \) and each \( (k-1) \)-capped labelled \( k \)-box is some scalar times identity, then the planar algebra is an Exchange Relation Planar Algebra in the sense of [Lan]. We will discuss some other examples of higher exchange relation planar algebra in the following section.

### 4. Subfactors with depth 2.

Let \( N \subset M \) be a finite index inclusion of \( \Pi_1 \) factors with depth 2 and \( N \subset M \subset M_1 \subset M_2 \subset \ldots \) be the Jones tower of basic construction, where \( e_k \) denotes the Jones projection in \( M_k \) which implements the conditional expectation of \( M_{k-1} \) onto \( M_{k-2} \). Therefore, \( N' \cap M_{k-1} \subset N' \cap M_{k} \subset e_{k+1} N' \cap M_{k+1} \) is a basic construction of finite dimensional \( C^* \)-algebras for all \( k \geq 1 \).

It follows from [Jon2] that there is a planar algebra, \( P_{N \subset M} \), associated to \( N \subset M \) where \( P_{N \subset M}^k = N' \cap M_{k-1} \) for all \( k \geq 0 \). For this planar
algebra, we may choose the generating set \( L = L_2 \) as the union of systems of matrix units of central summands of the ‘multi-matrix algebra’ \( P_2^{N \subset M} \); and we may take the set of relations as \( R = \{ X \in P(L) : \Phi^{N \subset M}(X) = 0 \} \) where \( \Phi^{N \subset M} \) is a presenting map for \( P^{N \subset M} \). (The reason we may make this choice of \( L \) is that \( P_{N \subset M} \) is a basic construction for all \( k \geq 0 \), so that \( P_{N \subset M} \) as an algebra is generated by \( P_{2 \subset M} \) and \( e_1, e_2, e_3, \ldots, e_{k-1} \).)

Let \( \Phi_R \) denote the quotient map from \( P(L) \) onto \( P_{\langle L, R \rangle} \). The inclusion \( L \subset P^{N \subset M} \) is seen to induce a planar algebra homomorphism \( \Gamma \) from \( P_{\langle L, R \rangle} \) into \( P_{N \subset M} \) such that the following diagram commutes:

\[
P(L) \xrightarrow{\Phi_R} P_{\langle L, R \rangle} \xrightarrow{\Gamma} P^{N \subset M}
\]

From the argument of the (parenthetical statement of the) previous paragraph, and the commutativity of the above diagram, it is seen that \( \Gamma \) is surjective. Now for \( X \in P(L) \), if \( \Gamma \circ \Phi_R(X) = 0 \), then \( \Phi^{N \subset M}(X) = 0 \) which implies \( X \in R \), that is, \( \Phi_R(X) = 0 \). Hence \( \Gamma \) is 1-1 and thereby a planar algebra isomorphism.

So we can view \( P (= P^{N \subset M}) \) as a planar algebra presented on \( L = L_2 \) in the above manner, with \( |L| < \infty \).

**Proposition 8.** Any planar algebra associated to a finite index inclusion of \( \Pi_1 \)-factors with depth 2 is a higher exchange relation planar algebra.

**Proof.** Let \( N \subset M \) be a finite index inclusion of \( \Pi_1 \)-factors with depth 2. To it, we associate the planar algebra, \( P_{\langle L, R \rangle} \) described above. We proceed now to systematically verify that \( P_{\langle L, R \rangle} \) indeed satisfies the several requirements for being a higher exchange relation algebra.

(o) If \( \gamma \) is a contractible loop, then

\[
\Phi^{N \subset M}(\gamma - \delta 1) = 0 \Rightarrow (\gamma - \delta 1) \in R \quad \text{for} \quad \delta = [M : N]^{\frac{1}{2}}.
\]

(i) A 2-capped \( L \)-labelled 2-box can — depending on the location of the marked point labelled 1 — be viewed as an element of \( P_0(L, R) \) or of \( P_{1,1}(L, R) \). But \( P_0(L, R) = \mathbb{C} = P_{1,1}(L, R) \). So, for any \( l \in L \), it follows that \( R \) contains \( l^{(2)} \) — (some constant)1.

(ii) Since the conditional expectation of \( P_2(L, R) = N' \cap M_1 \) onto \( P_1(L, R) = N' \cap M \) is given by a scalar multiple of an appropriate 1-capping of 2-tangles in \( P_2(L, R) \) and since \( P_2(L, R) \) is linearly spanned by \( L \)-labelled 2-boxes, we see that \( P_1(L, R) \) is linearly spanned by \( \{ \Phi_R(l^{(1)}) : l \in L \} \). In particular, \( \text{span} \{ \Phi_R(l^{(1)}) : l \in L \} = P_1(L, R) \) is closed under multiplication.
(iii) The depth 2 assumption says that \( N' \cap M \subset N' \cap M_1 \subset N' \cap M_2 \) is a basic construction; so \( P_3(L, R) = P_2(L, R)E_2P_2(L, R) \) (where \( E_2 \) denotes the usual ‘conditional expectation tangle — see [Jon2]). Again, \( P_2(L, R) \) is the span of \( L \)-labelled 2-boxes (since \( P(L) \) is, by choice of \( L \)). Hence for \( A, B \in L \), we find that

is a linear combination of

for \( C, D \in L \) (* for the internal boxes with labels, \( A, B \) can be arbitrary as long as the tangle makes sense). This shows that \( \mathcal{R} \) satisfies Criterion (iii) in the definition of higher exchange relation planar algebras.

Thus, \( P(L, \mathcal{R}) \) is a higher exchange relation planar algebra. \( \square \)

5. \( E_6 \) and \( E_8 \) subfactors.

We consider the subfactors with principal graphs, \( E_6 \) and \( E_8 \), given in Figure 10.

![Figure 10](image)

Our goal is to show that the planar algebras associated to these subfactors are of the higher exchange relation type. We will be using facts stated in [Jon4]. In fact, many of the arguments here, particularly those concerned with the rotation, are from [Jon4].
Let $P^6$ (resp. $P^8$) be the planar algebra associated to the subfactor with principal graph $E_6$ (resp. $E_8$). Since $E_6$ (resp. $E_8$) has a unique vertex of degree 3, which is at distance 2 (resp. 4) from $\ast$, we see from [LaSu] that $P^6$ (resp. $P^8$) is generated by one 3-box, $R^6$ (resp. one 5-box $R^8$) which can be taken to be any element of $P_3^6 \setminus TL_3$ (resp. $P_5^8 \setminus TL_5$). We assume $R^6$ (resp. $R^8$) is a unit vector orthogonal to $TL$. Since $TL$ is invariant under rotation — see ([Jon2]) — it follows that $R^6$ (resp. $R^8$) is an eigenvector for $\rho_3$ (resp. $\rho_5$) corresponding to an eigenvalue of modulus one — in view of the ‘orthogonality’ preserving nature of the rotation (see Remark 2(d), and [Jon2]).

In general, any planar algebra satisfies Conditions (o) and (i) in the definition of higher exchange relation planar algebra. So both $P^6$ and $P^8$ satisfy Conditions (o) and (i).

Next we note that our choice of $R^6$ (resp. $R^8$) — as an element in the orthogonal complement of the Temperley-Lieb subalgebra — implies that the result of any 1-capping of $R^6$ (resp. $R^8$) yields zero and consequently any 2-capping of $R^6$ (2-capping or 4-capping of $R^8$) is zero in $P^6$ (resp. $P^8$). Thus Condition (ii) is satisfied.

Now we shall see that both $P^6$ and $P^8$ satisfy Condition (iii) in the definition of a higher exchange relation planar algebra, but for different reasons.

For $P^6$, we have $L = L_3 = \{R^6\}$. In Condition (iii) the picture, denoted by $A$, is a 5-tangle. So we consider $P^6_5$. The dimension of $P^6_5$ can be calculated from the principal graph and turns out to be 77.

![Figure 11](image)

Next we consider the dimension of the subspace spanned by 5-tangles with exactly one 3-box with label $R^6$. Actually there are 45 such tangles but not all of them are linearly independent. Jones [Jon4] deduces from the work [GrLe] that the dimension of the subspace spanned by these 5-tangles is
35. This subspace lies in the orthogonal complement of the Temperley-Lieb subalgebra. We know that \( \dim(TL_5) = 42 \). Thus we see that \( \dim(P^6_5) = 77 = 35 + 42 \). Therefore, \( P^6_5 \) is generated by 5-tangles with at most one 3-box. In particular, our 5-tangle \( A \) of Condition (iii) is a linear combination of 5-tangles with at most one internal 3-box. So \( P^6 \) satisfies Condition (iii), and is hence a higher exchange relation planar algebra.

For \( P^8 \), the tangle \( A \) of Condition (iii) is a 9-tangle. So we consider \( P^8_9 \). We first look at the subspace generated by 9-tangles with exactly one 5-box. The dimension of this subspace was found (in [Jon4], using [GrLe]) to be 2244. This subspace is also orthogonal to the Temperley-Lieb subalgebra of \( P^8_9 \); and we have \( \dim(TL_9) = 4862 \). Using the principal graph, we can calculate and see that \( \dim(P^8_9) = 7107 = 2244 + 4862 + 1 \).

![Figure 12.](image)

Next consider the subspace — call it \( W \) — which is generated by 9-tangles with at most one labelled 5-box. So we have shown that \( W \) has codimension 1 in \( P^8_9 \). Now consider the 9-tangle in Figure 13 (where we write \( R \) for \( R^8 \)).

Let us denote the above tangle by \( A_1 \). Note that the *’s for the internal boxes in \( A_1 \) are not important because \( R^8 \) forms an eigenvector for the rotation, \( \rho_9 \). If \( A_1 \) belongs to \( W \), then we are through. If not, then \( P^8_9 \) is linearly spanned by \( W \cup \{A_1\} \). It is enough, by Remark 4(b), to verify Condition (iii) for the 9-tangle in Figure 14, which will be denoted by \( A_2 \).
Figure 13.

Figure 14.

Being an element of $P_9^8$, $A_2$ is a linear combination of $A_1$ and elements of $W$. So, $P_9^8$ satisfies Condition (iii), and is, indeed, a higher exchange relation planar algebra.

Acknowledgement. I would like to thank V.S. Sunder for valuable discussions and Zeph Landau for allowing me access to his preprint.

References


Received January 18, 2002 and revised July 4, 2002.

INSTITUTE OF MATHEMATICAL SCIENCES
CHENNAI 600113
INDIA
E-mail address: shami@imsc.ernet.in
KLEIN BOTTLE SURGERY AND GENERA OF KNOTS

Kazuhiro Ichihara and Masakazu Teragaito

In this paper, we study the creation of Klein bottles by surgery on knots in the 3-sphere. For non-cabled knots, it is known that the slope corresponding to such surgery is an integer. We give an upper bound for the slopes yielding Klein bottles in terms of the genera of knots.

1. Introduction.

In this paper, we will study the creation of Klein bottles by surgery on knots in the 3-sphere $S^3$. Let $K$ be a knot in $S^3$, and let $E(K)$ be its exterior. A slope on $\partial E(K)$ is the isotopy class of an essential simple closed curve in $\partial E(K)$. As usual, the slopes on $\partial E(K)$ are parameterized by $\mathbb{Q} \cup \{1/0\}$, where $1/0$ corresponds to a meridian slope (see [R]). For a slope $r$, $K(r)$ denotes the closed 3-manifold obtained by $r$-Dehn surgery on $K$. That is, $K(r) = E(K) \cup V$, where $V$ is a solid torus glued to $E(K)$ along their boundaries in such a way that $r$ bounds a meridian disk in $V$.

Suppose that $K(r)$ contains a Klein bottle. Then $K(r)$ is shown to be reducible, toroidal or Seifert-fibered [L], and therefore it is non-hyperbolic. Gordon and Luecke [GL] showed that such a slope $r$ is integral when $K$ is hyperbolic. Furthermore, such a slope must be divisible by four in this case [T1]. These results together with the bound on exceptional surgeries [A, Theorem 8.1] imply that there are at most three surgeries creating Klein bottles on a hyperbolic knot in $S^3$.

However, unfortunately, there is no universal upper bound on the absolute values of such slopes. That is, for any positive number $N$, there exists a hyperbolic knot in $S^3$ which admits $r$-surgery creating a Klein bottle for $r > N$. See Section 5.

In [T1], we gave an upper bound on the absolute value of such a slope $r$ in terms of the genera of knots. That is, for a non-cabled knot $K$, $|r| \leq 12g(K) - 8$, where $g(K)$ is the genus of $K$. Indeed, we had a better inequality $|r| \leq 8g(K) - 4$ if $r$ is not the boundary slope of a once-punctured Klein bottle spanned by $K$.

The main theorem of this paper greatly improves both estimations:
Theorem 1.1. Let $K$ be a non-cabled knot in $S^3$. If $K(r)$ contains a Klein bottle, then $|r| \leq 4g(K) + 4$. Moreover, if $r$ is not the boundary slope of a once-punctured Klein bottle spanned by $K$, then $|r| \leq 4g(K) - 4$.

We remark that such a slope can be non-integral for a cable knot. In fact, $16/3$-surgery on the right-handed trefoil yields a prism manifold which contains a Klein bottle. Also we remark that, as far as we know, there is no example of the case that $r$ is not the boundary slope of a once-punctured Klein bottle spanned by $K$. (The knots of [BH, Propositions 18,19] are strong candidates.)

The extremal case $|r| = 4g(K) + 4$ can be described completely in the following:

Theorem 1.2. Let $K$ be a non-cabled knot in $S^3$. Suppose that $K(r)$ contains a Klein bottle. If $|r| = 4g(K) + 4$, then $K$ is the connected sum of the $(2,m)$-torus knot and the $(2,n)$-torus knot, and $r = 2m + 2n$, where $m, n \neq \pm 1$ are odd integers with the same sign.

Corollary 1.3. Let $K$ be a hyperbolic knot in $S^3$. If $K(r)$ contains a Klein bottle, then $|r| \leq 4g(K)$. Moreover, if $|r| = 4g(K)$, then $K$ bounds a once-punctured Klein bottle whose boundary slope is $r$.

For example, $\pm 4$-surgery on the figure eight knot yield Klein bottles. Clearly, each slope bounds a once-punctured Klein bottle. (Consider a checkerboard surface of its standard diagram.) Since it has genus one, the above estimation is sharp. In Section 5, such a hyperbolic knot will be given for each genus.

The authors are grateful to the referee for many useful comments. We also thank Seungsang Oh for Lemma 3.6.

2. Preliminaries.

Throughout the paper, $K$ is assumed to be a non-cabled knot. We denote by $g$ the genus of $K$. Suppose that $K(r)$ contains a Klein bottle $\hat{P}$ for a slope $r$. In general, 0-surgery can yield a Klein bottle, but we may assume $r \neq 0$ to prove Theorem 1.1. Thus we may assume $r > 0$. Let $k$ be the core of the attached solid torus $V$. We may assume that $k$ intersects $\hat{P}$ transversely, and that $\hat{P}$ is chosen to minimize $p = |\hat{P} \cap k|$ among all Klein bottles in $K(r)$. Then $P = \hat{P} \cap E(K)$ is a punctured Klein bottle properly embedded in $E(K)$ with $|\partial P| = p$. We note that $p \geq 1$ and $p$ is odd. Otherwise a closed non-orientable surface can be obtained by attaching suitable annuli along $\partial P$.

Lemma 2.1. $P$ is incompressible and boundary incompressible in $E(K)$.

Proof. See Lemmas 2.1 and 2.2 of [T1]. □
Lemma 2.2. \( r \) is an integer divisible by four.

Proof. See Lemmas 2.3 and 2.4 of [T1].

Let \( Q \subset E(K) \) be a minimal genus Seifert surface of \( K \). Then \( Q \) is incompressible and boundary incompressible in \( E(K) \). Let \( \hat{Q} \) denote the closed surface obtained by capping \( \partial Q \) off by a disk. We may assume that \( P \) and \( Q \) intersect transversely, and that \( P \cap Q \) contains no circle component which bounds a disk in \( P \) or \( Q \) by the incompressibility of these surfaces. Also, we can assume that each component of \( \partial P \) intersects \( \partial Q \) in exactly \( r \) points, since \( \partial Q \) has slope 0.

Let \( G_Q \) be the graph in \( \hat{Q} \) obtained by taking as the fat vertex the disk \( \hat{Q} - \text{Int}Q \) and as edges the arc components of \( P \cap Q \). Similarly, \( G_P \) is the graph in \( \hat{P} \) whose vertices are the \( p \) disks \( \hat{P} - \text{Int}P \) and whose edges are the arc components of \( P \cap Q \). Thus the edges of \( G_P \) and \( G_Q \) are in one-one correspondence. When \( p > 1 \), number the components of \( \partial P \), 1, 2, \ldots, \( p \) in sequence along \( \partial E(K) \). This induces a numbering of the vertices of \( G_P \). Each endpoint of an edge in \( G_Q \) has a label, namely the number of the corresponding component of \( \partial P \). Thus the labels 1, 2, \ldots, \( p \) appear in order around the vertex of \( G_Q \) repeated \( r \) times. An edge with labels \( i \) and \( j \) at its endpoints is called a \((i,j)\)-edge. If an edge has a label \( i \) at least one endpoint, it is called an \( i \)-edge. If both endpoints have label \( i \), then it is called a level \( i \)-edge, or simply a level edge. Since \( G_Q \) has just one vertex, the edges of \( G_P \) have no labels. A trivial loop in a graph is a length one cycle which bounds a disk face of the graph.

Lemma 2.3. Neither \( G_P \) nor \( G_Q \) contains trivial loops.

Proof. This is Lemma 3.1 in [T1].

Although \( P \) is non-orientable, we can establish a parity rule as a natural generalization of the usual one [CGLS]. Here we use a restricted form, because one graph has just one vertex. First orient all components of \( \partial P \) so that they are mutually homologous on \( \partial E(K) \). Also consider an orientation to \( \partial Q \). Let \( e \) be an edge of \( G_P \) (and \( G_Q \) simultaneously). Let \( D \) be a regular neighborhood of \( e \) on \( P \). Then \( D \) is a disk, and \( \partial D = a \cup b \cup c \cup d \), where \( a \) and \( c \) are arcs in \( \partial P \) with induced orientations from \( \partial P \). If \( a \) and \( c \) have the same direction along \( \partial D \), then \( e \) is said to be positive in \( G_P \), negative otherwise. See Figure 1. Similarly we define positive and negative edges in \( G_Q \). Since \( \partial E(K) \) is a torus and \( E(K) \) is orientable, we have the following expression of the parity rule:

Lemma 2.4 (Parity rule). Each edge of \( G_Q \) (\( G_P \), resp.) is positive (negative, resp.).

Throughout this section, we assume $p > 1$. This means that $r$ is not the boundary slope of a once-punctured Klein bottle spanned by $K$.

A sequence of edges in $G_Q$ is called a cycle. Since $G_Q$ has a single vertex, this is not a cycle in a sense of graph theory. Let $D$ be a disk face of $G_Q$. Then $\partial D$ is an alternating sequence of edges and corners (subarcs of $\partial Q$). Thus we can regard that $\partial D$ defines a cycle. If $D$ is bounded by only $i$-edges, and all the $i$-edges have the same pair of labels $\{i, i+1\}$ at their endpoints, then the cycle defined by $\partial D$ is called a Scharlemann cycle with the label pair $\{i, i+1\}$. The number of edges in a Scharlemann cycle $\sigma$ is defined to be the length of $\sigma$. In particular, a Scharlemann cycle of length two is called an $S$-cycle. A triple of successive parallel edges $\{e_{-1}, e_0, e_1\}$ is called a generalized $S$-cycle if $e_0$ is a level $i$-edge and both $e_{-1}$ and $e_1$ are $(i-1, i+1)$-edges.

Lemma 3.1.  
(i) $G_Q$ does not contain a Scharlemann cycle.  
(ii) $G_Q$ does not contain a generalized $S$-cycle.

Proof. These are Lemmas 3.2 and 3.3 in [T1]. (In fact, [T1, Lemma 3.2] treats only an $S$-cycle, but the argument works in general.) □

Lemma 3.2. At most two vertices of $G_P$ are incident to negative loops.

Proof. Let $e$ be a negative loop at a vertex $v$ in $G_P$. Then $N(v \cup e)$ is a Möbius band in $\hat{P}$. Since a Klein bottle contains at most two disjoint Möbius bands, the result follows. □

By Lemma 3.2, there are at most two vertices $u$ and $v$ of $G_P$ which are incident to negative loops. This means that $G_Q$ has at most two kinds of level edges. These are called level $u$-edges and level $v$-edges.

Lemma 3.3. There are at most $r/2$ level $u$ ($v$)-edges in $G_Q$. 

Figure 1.
Proof. Since $u$ has degree $r$ in $G_P$, there are at most $r/2$ negative loops at $u$. The result follows from the parity rule.

Since $p \geq 3$, we can choose a vertex $x$ of $G_P$ which is not incident to a negative loop by Lemma 3.2. We fix this $x$ hereafter. Let $\Gamma_x$ be the subgraph consisting of all $x$-edges and the vertex of $G_Q$. Since $G_Q$ does not contain level $x$-edges, $\Gamma_x$ has just $r$ edges. A disk face of $\Gamma_x$ is called an $x$-face.

**Lemma 3.4.** Any $x$-face contains at least one level edge of $G_Q$ in its interior.

**Proof.** Assume that an $x$-face $D$ does not contain a level edge. Then $D$ contains a Scharlemann cycle by [HM, Lemma 5.2]. This contradicts Lemma 3.1.

**Lemma 3.5.** If $r > 4g - 4$, then there are two $x$-faces $D_u$ and $D_v$ such that $D_u$ contains only level $u$-edges and $D_v$ contains only level $v$-edges.

**Proof.** Let $X$ be the number of $x$-faces. Then an Euler characteristic calculation for $\Gamma_x$ gives

$$1 - r + \sum_{f: \text{a face of } \Gamma_x} \chi(f) = 2 - 2g.$$ 

Thus $X \geq \sum \chi(f) = 1 - 2g + r$. Since $r$ is divisible by four, $r \geq 4g$. Then $X \geq r/2 + 1$. The result follows from Lemmas 3.3 and 3.4.

We show that the existence of the $x$-faces $D_u, D_v$ gives a contradiction. Let $D_\alpha = D_u$ or $D_v$. Thus $D_\alpha$ contains only level $\alpha$-edges.

Let $D$ be a disk face of $G_Q$. Recall that $\partial D$ is an alternating sequence of edges of $G_Q$ and corners. A corner with labels $\{i, i+1\}$ at its endpoints is denoted by $(i, i+1)$. If $\partial D$ contains only two kinds of corners $(\alpha, \alpha + 1)$ and $(\alpha - 1, \alpha)$, then $D$ is called a two-cornered face. Such a notion was first used in [H].

**Lemma 3.6.** $D_\alpha$ contains a pair of two-cornered faces sharing a level $\alpha$-edge on their boundaries, such that at least one of such two-cornered faces contains only one level $\alpha$-edge.

**Proof.** If $D_\alpha$ has no non-level diagonal edge (that is, each edge in $D_\alpha$ joining non-adjacent corners along $\partial D_\alpha$ is level), then set $E = D_\alpha$. Suppose that $D_\alpha$ contains a diagonal edge $e$ which has distinct labels $\{a, b\}$. Note that $a \neq x, b \neq x$. Without loss of generality, assume that the labels appear in counterclockwise order around the corners of $\partial D_\alpha$, and that $a < b < x$. This means that these labels $a, b, x$ appear in this order around the vertex of $G_Q$. (Thus three inequalities $a < b < x, b < x < a$ and $x < a < b$ are equivalent.) Formally, we construct a new $x$-face $D'$ as follows: Consider that $e$ is oriented from the endpoint with label $a$ to the other. Discard the
half (disk) of $D_\alpha$ right to $e$. Insert additional edges to the right of $e$ and parallel to $e$ until the label $x$ first appear at one end of this parallel family of edges. Possibly, the last edge has label $x$ at its both endpoints. But, except the last edge, there is no level edge among the additional edges. In fact, the label sets $I, J$ indicated in Figure 2 are disjoint, except the case where the last edge is a level $x$-edge. Let $D'$ be the union of the left side of $e$ and the bigons among these parallel family. Then $D'$ is an $x$-face. See Figure 2. There is no two-cornered face among the additional bigons. Repeat this process for every diagonal edge in $D'$ which is not a level $\alpha$-edge, then we finally get a new $x$-face $E$.

Thus all diagonal edges in $E$ are level $\alpha$-edges, and $\partial E$ consists of $x$-edges. As remarked before, there may be level $x$-edges in $\partial E$. If $E$ does not contain a level $\alpha$-edge, then [HM, Lemma 5.2] says that there is a Scharlemann cycle $\sigma$ in $E$. By the construction of $E$, $\sigma$ lies in $D_\alpha$, and that is, $\sigma$ lies in $G_Q$. But this is impossible by Lemma 3.1. Hence $E$ contains a level $\alpha$-edge.

**Claim 3.7.** Any face adjacent to a level $\alpha$-edge in $E$ is two-cornered.

**Proof.** Let $e$ be a level $\alpha$-edge in $E$, and let $f$ be a face adjacent to $e$. Note that $\partial f$ may contain other level $\alpha$-edges. Let $(a_i, a_i + 1)$ ($i = 1, 2, \ldots, n$) be the corners on $\partial f$ between successive level $\alpha$-edges (possibly, the same one) on $\partial f$, which appear in order around $f$ when we go around clockwise. Thus $a_1 = \alpha$ and $a_n = \alpha - 1$. See Figure 3.

Let $e_i$ be the edge on $\partial f$ connecting the points with labels $a_i + 1$ and $a_{i+1}$ for $i = 1, 2, \ldots, n - 1$. Note that $e_i$ is neither a level $x$-edge nor a diagonal edge in $E$. Also, $e_i$ can be an $x$-edge, otherwise it is parallel to an $x$-edge (which may be level).

First consider $e_{n-1}$. If $e_{n-1}$ is an $x$-edge, then $x = a_n < a_{n-1} + 1$ or $a_n < a_{n-1} + 1 = x$ (of course, for any two labels $a, b$, the inequalities $a < b$
and $b < a$ are equivalent). Hence $x \leq a_n < a_{n-1} + 1 \leq a$
More precisely, when we go around the vertex of $G_Q$ (in counterclockwise
direction), the two labels $a_n, a_{n-1} + 1$ appear in this order between the
successive $x$'s. If $e_{n-1}$ is not an $x$-edge, then it is parallel to an $x$-edge $e'$
(possibly, a level $x$-edge) in $\partial E$ by the construction of $E$. Let $F$
be the family of mutually parallel edges containing $e_{n-1}$ and $e'$. We refer to the
end of $F$ containing the end point of $e_{n-1}$ with label $a_{n-1} + 1$ as the left
end. Assume that the label $x$ appears at the left end of $F$ (in fact, at the “left
end” of $e'$). By Lemma 3.1 and the construction of $E$, the label $a_n$
cannot appear at the left end of $F$. Hence three labels $a_n, a_{n-1} + 1, x$
appear in this order, that is, $a_n < a_{n-1} + 1 < x$. If $x$ appears at the right end of $F$, then we have the same inequality similarly. Hence $x \leq a_n < a_{n-1} + 1 \leq x$ holds again. Thus we always have $x \leq a_n \leq a_{n-1} < a$
Next, consider the edge $e_{n-2}$. By the same argument as above, we have
$x \leq a_{n-1} \leq a_{n-2} < x$. Continuing in this way, we eventually get $x \leq a_n \leq a_{n-1} \leq \cdots \leq a_1 < x$. This means that the labels $a_n, a_{n-1}, \ldots, a_2, a_1$
in this order between the successive $x$'s. But recall that $a_n = \alpha - 1$ and
$a_1 = \alpha$ are successive. Hence $a_1 = \cdots = a_j = \alpha$ and $a_{j+1} = \cdots = a_n = \alpha - 1$
for some $j$. Thus we have proved that $f$ is two-cornered face.

Choose a level $\alpha$-edge $e$ in $E$, which is outermost among level $\alpha$-edges
in $E$. That is, $e$ cuts a disk $E'$ off from $E$ which contains no level $\alpha$-edge
except $e$. Let $f$ and $g$ be the faces adjacent to $e$. Then these are a desired
pair of two-cornered faces. Note that one of them can be a bigon, but then
the other has at least three sides, since $G_Q$ does not contain a generalized
$S$-cycle by Lemma 3.1.

Let $\hat{T}$ be the torus which is the boundary of a thin regular neighborhood
$N(\hat{P})$ of $\hat{P}$, and let $T$ be the intersection of $\hat{T}$ with $E(K)$. Then $T \cap Q$
gives rise to a pair of graphs $\{G_T, G_T^Q\}$, where $G_T$ is a “double cover” of
$G_P$, and each edge of $G_Q$ corresponds to a bigon of $G_T^Q$. Let $i_1, i_2$
be the vertices of $G_T$ corresponding to $i$-th vertex of $G_P$ such that they appear in
the same order as the vertices of $G_P$. Thus $11, 12, 21, 22, \ldots, p1, p2$ appear along $\partial E(K)$ in this order. In particular, each level $u$-edge ($v$-edge, resp.) of $G_Q$ yields an S-cycle in $G^T_Q$ with label-pair $\{u1, u2\}$ ($\{v1, v2\}$, resp.).

By Lemma 3.6, $D_\alpha$ contains a pair of two-cornered faces sharing a level $\alpha$-edge. These give an S-cycle $\sigma_\alpha$ and two faces $f_\alpha$, $g_\alpha$ adjacent to $\sigma_\alpha$. Note that $f_\alpha$ and $g_\alpha$ contain only $(\alpha2, (\alpha + 1)1)$ and $((\alpha - 1)2, \alpha1)$ corners. By Lemma 3.6, we may assume that $f_\alpha$ contains only one $(\alpha1, \alpha2)$-edge, which is an edge of $\sigma_\alpha$. Also remark that $f_\alpha$ contains only one $((\alpha + 1)1, (\alpha - 1)2)$-edge. By the construction of $\hat{T}$, there are disjoint annuli in $\hat{T}$, which contain the edges of $\sigma_u$ and $\sigma_v$, respectively. Note that the centerlines of these annuli are essential on $\hat{T}$.

**Lemma 3.8.** $u$ and $v$ are not adjacent on $\partial E(K)$.

*Proof.* Assume that $u$ and $v$ are adjacent. For simplicity, let $u = 2$ and $v = 3$.

Let $X$ be the number of $x$-faces. As in the proof of Lemma 3.5, $X \geq r/2 + 1$. Thus $r \leq 2X - 2$.

Let $X_2, X_3$ denote the number of $x$-faces which contain only level 2 or 3-edges respectively, and let $X_1$ be the number of $x$-faces containing both kinds of level edges. By Lemma 3.4, $X = X_1 + X_2 + X_3$. Count the number of occurrence of label 3 in $G_Q$. Each $x$-face containing only level 2-edges contains at least two occurrences of label 3. The other $x$-faces contain level 3-edges, which do not lie on the boundaries. Hence $2X_1 + 2X_2 + 2X_3 \leq r$, since each label appear $r$ times around the vertex of $G_Q$. Thus $2X_1 + 2X_2 + 2X_3 \leq r \leq 2X - 2 = 2(X_1 + X_2 + X_3) - 2$, which is a contradiction. \qed

Let $A_u$ ($A_v$, resp.) be the subgraph of $G_T$ consisting of four vertices $u1, u2, (u - 1)2, (u + 1)1$ ($v1, v2, (v - 1)2, (v + 1)1$, resp.) and the edges of $\sigma_u$, $\partial f_u$ and $\partial g_u$ ($\sigma_v, \partial f_v, \partial g_v$, resp.). As noted in the proof of Lemma 3.6, $g_u$ and $g_v$ have at least three sides, and hence $A_u$ and $A_v$ are connected. Hence there is an annulus $A_u$ ($A_v$, resp.) in $\hat{T}$ which contains the edges of $\sigma_u$, $\partial f_u$ and $\partial g_u$ ($\sigma_v, \partial f_v, \partial g_v$, resp.). By Lemma 3.8, the vertices $u1, u2, v1, v2, (u - 1)2, (u + 1)1, (v - 1)2, (v + 1)1$ are distinct. We may assume that $A_u$ and $A_v$ are disjoint, and that one boundary component of $A_u$ ($A_v$, resp.) is very near to the edges of $\sigma_u$ ($\sigma_v$, resp.). Moreover, these boundary components bound M"obius bands $M_u$ and $M_v$, respectively, in $N(\hat{P})$ meeting the core of the attached solid torus $V$ in a single point. (Consider the union of the bigon face of $\sigma_\alpha$ and the 1-handle $H$ bounded by the vertices $\alpha1, \alpha2$. By shrinking $H$ radially to its core, we obtain a M"obius band, and then perturb it to be transverse to the core of $V$.)

Let $q_u$ be the number of vertices contained in $A_u$. Let $H_1$ and $H_2$ are disjoint 1-handles on $V$ bounded by the vertices of $(u - 1)2$ and $u1, u2$ and $(u + 1)1$. Consider $W_u = N(A_u \cup H_1 \cup H_2 \cup f_u \cup g_u) \subset K(r) - \text{Int} N(\hat{P})$. 

324 K. ICHIHARA AND M. TERAGAITO
Lemma 3.9. $\partial W_u$ consists of one or two tori.

Proof. Let $W'_u = N(A_u \cup H_1 \cup H_2)$. Then $W'_u$ is a handlebody of genus three. Since $\partial f_u$ is non-separating on $\partial W'_u$, attaching a 2-handle $N(f_u)$ gives a genus two surface from $\partial W'_u$.

We claim that $\partial f_u$ and $\partial g_u$ are not parallel on $\partial W'_u$. If $\partial f_u$ and $\partial g_u$ are parallel on $\partial W'_u$, then these represent the same element of $\pi_1(W'_u)$. Taking as a base "point" a subdisk of $A_u$ as shown in Figure 4, we have $\pi_1(W'_u) = \langle x_1, x_2, y \rangle$, where $x_1$ ($x_2$, resp.) is represented by a core of $H_1$ ($H_2$, resp.) going from vertex $(u-1)2$ ($u2$, resp.) to vertex $u1$ ($(u+1)1$, resp.), and $y$ is represented by the edge of $\sigma_u$ not in the base "point" going from vertex $u1$ to vertex $u2$. We may assume that the $(u1,u2)$-edge on $\partial f_u$ is contained in the base "point". Then $\partial f_u$ never contain $x_1yx_2$, although $\partial g_u$ contains it (in the appropriate directions). This is because $\partial f_u$ contains just one $(u1,u2)$-edge. Therefore $\partial f_u$ and $\partial g_u$ are not parallel on $\partial W'_u$, and then $\partial W_u$ cannot be a union of a 2-sphere and a genus two surface. Thus $\partial W_u$ is a torus or two tori according to whether the attaching sphere of $N(g_u)$ is non-separating or separating on $\partial(W'_u \cup N(f_u))$. □

![Figure 4.](image_url)

Then $\partial W_u$ contains a torus $F$ containing $A_u$, since $W_u \cap \hat{T} = A_u$. Note that the core of $A_u$ is essential on $F$. If not, one component of $\partial A_u$ bounds a disk $D$ in $K(r) - N(\hat{P})$. Then $M_u \cup D$ or $M_u \cup A_u \cup D$ is a projective plane in $K(r)$. Take a parallel copy $D'$ of $D$ in $K(r) - N(\hat{P})$, so that $D'$ is disjoint from $A_u$. Then the union of $D'$, $M_v$ and the annulus on $\hat{T}$ bounded by $\partial D'$ and $\partial M_v$, which is disjoint from $A_u$, forms another projective plane in $K(r)$. Thus $K(r)$ contains two disjoint projective planes, but this is impossible, because $H_1(K(r))$ would be non-cyclic.
Let $A'_u$ be the remaining annulus of the torus. By the construction, $A'_u$ meets the core of $V$ in at most $q_u - 4$ points. Similarly, we obtain an annulus $A'_v$ by using $f_v, g_v$.

The edges of $\sigma_u$ and $\sigma_v$ separate $\hat{T}$ into two annuli $B_1$ and $B_2$. Each $\text{Int}B_i$ contains $p - 2$ vertices, since $G_T$ is a double cover of $G_P$. We may assume that the edges of $\partial f_u$ and $\partial g_u$ are contained in $B_1$.

First assume that the edges of $\partial f_v, \partial g_v$ are contained in $B_1$. Let $B'_1 \subset B_1$ be the annulus region bounded by $\partial A'_u$ and $\partial A'_v$. Then the union $M_u \cup A'_u \cup B'_1 \cup M_v$ gives a new Klein bottle in $K(r)$, which meets the core of $V$ in at most $p - 4$ points. This contradicts the minimality of $p$.

Next assume that the edges of $\partial f_v, \partial g_v$ are contained in $B_2$. Let $B''_1 \subset B_1$ be the annulus region bounded by $\partial A'_u$ and $\partial A'_v$. Then the union $M_u \cup A'_u \cup B''_1 \cup M_v$ gives a new Klein bottle in $K(r)$, which meets the core of $V$ in at most $p - 2$ points, a contradiction.

Thus we have proved Theorem 1.1 when $p > 1$.

4. Special case.

In this section, we consider the case where $p = 1$. Assign the points of $\partial P \cap \partial Q$ the labels $1, 2, \ldots, r$ along $\partial Q$ sequentially. Then the labels are also sequential along $\partial P$, since $r$ is integral.

**Lemma 4.1.** If $G_Q$ has parallel edges, then $r = 4$.

*Proof.* This is Lemma 4.2 in [T1].

Thus we may assume that $G_Q$ has no parallel edges hereafter.

**Lemma 4.2.** If two edges of $G_P$ are parallel, then their endpoints appear alternately around the vertex of $G_P$.

*Proof.* This follows from that all edges of $G_P$ are negative.

**Lemma 4.3.** Suppose that $G_Q$ contains a separating edge $e$. Then one component of $Q - e$ contains no edge of $G_Q$.

*Proof.* Assume that each component of $Q - e$ contains an edge $e_1$ and $e_2$ respectively. Since $G_P$ consists of at most two parallel families of (negative) edges (cf. [T1, Section 4]), some two of $e, e_1, e_2$ are parallel in $G_P$. But this is impossible by Lemma 4.2.

**Lemma 4.4.** If $G_Q$ contains a separating edge, then $r \leq 4g$.

*Proof.* Let $e$ be a separating edge in $G_Q$, and let $Q_1$ and $Q_2$ be the components of $Q - e$. By Lemma 4.3, we may assume that $Q_2$ contains no edge. If $Q_1$ contains a separating edge $e_1$, then $e$ and $e_1$ are not parallel in $G_P$ by Lemma 4.2. If $r > 4$, then $G_Q$ contains another edge $e_2$, which is parallel to $e$ or $e_1$ in $G_P$. But Lemma 4.2 gives a contradiction. Hence we may assume...
that $Q_1$ contains no separating edges. In fact, no edge in $Q_1$ is parallel to $e$ in $G_P$ by Lemma 4.2. Thus $G_P$ consists of $e$ and a parallel family of $r/2 - 1$ edges. By examining the labels of edges, we see that $G_Q$ has just three faces.

For $\hat{Q}$, we have

$$1 - \frac{r}{2} + \sum_{f: \text{a face of } G_Q} \chi(f) = 2 - 2g.$$ 

Thus $\sum \chi(f) = 1 - 2g + r/2$. Here $\sum \chi(f) = \sum_{f \neq Q_2} \chi(f) + \chi(Q_2) \leq \sum_{f \neq Q_2} \chi(f) - 1$. Thus $2 - 2g + r/2 \leq \sum_{f \neq Q_2} \chi(f)$. Since $G_Q$ has at most two disk faces, $2 - 2g + r/2 \leq 2$, and therefore $r \leq 4g$. □

**Lemma 4.5.** If $G_Q$ contains no separating edges, then $r \leq 4g + 4$.

**Proof.** Recall that $G_P$ consists of at most two families $A$ and $B$ of parallel edges. Let $|A|, |B|$ denote the number of edges in $A$ and $B$ respectively. Since $|A| + |B| = r/2$ is even, $|A|$ and $|B|$ have the same parity.

If $|A|$ and $|B|$ are even, then we see that $G_Q$ has just one face by examining the labels of edges. See Figure 5.

![Figure 5](image_url)

Then $1 - r/2 + \sum \chi(f) = 2 - 2g$, and thus $1 - 2g + r/2 = \sum \chi(f) \leq 1$. Therefore $r \leq 4g$.

If $|A|$ and $|B|$ are odd, then $G_Q$ has just three faces. Thus $1 - 2g + r/2 = \sum \chi(f) \leq 3$, and then $r \leq 4g + 4$. □
Lemmas 4.4 and 4.5 give the proof of Theorem 1.1 when $p = 1$. In fact, we can give the same upper bound $4g + 4$ by a 4-dimensional technique. We thank Seiichi Kamada for this suggestion. Consider $S^3 = \partial B^4$. The knot $K$ bounds $P$ and $Q$. Then pushing $P$ slightly into $B^4$ gives a closed non-orientable surface $P \cup Q$ embedded in $B^4$. Note $\chi(P \cup Q) = -2g$, where $g$ is the genus of $Q$. By Whitney-Massey Theorem (cf. [K]), the Euler number $e(P \cup Q)$ can vary between $2\chi(P \cup Q) - 4$ and $4 - 2\chi(P \cup Q)$. Thus $|e(P \cup Q)| \leq 4g + 4$. But $e(P \cup Q)$ is equal to the self-intersection number of $P \cup Q$, which is exactly the boundary slope of $P$ (see [K]).

5. Extremal case.

In this section, we examine the extremal case where $r = 4g + 4$, and prove Theorem 1.2. Recall that the points of $\partial P \cap \partial Q$ are labeled $1, 2, \ldots, r$ sequentially along $\partial P$ (and $\partial Q$) as in Section 4.

Assume $r = 4g + 4$. By the proof of Lemma 4.5, $G_P$ consists of two parallel families $A$ and $B$ such that $m = |A|$ and $n = |B|$ are odd. In fact, $|A|, |B| > 1$, otherwise $G_Q$ contains a trivial loop. Let $a_1, a_2, \ldots, a_m$ and $b_1, b_2, \ldots, b_n$ be the edges of $A$ and $B$ respectively, where they are numbered successively. That is, $a_i$ has labels $i$ and $i + m$, and $b_j$ has $2m + j$ and $2m + n + j$. See Figure 6, where the two end circles of the cylinder are identified with a suitable involution to form a Klein bottle $\hat{P}$.

![Figure 6](image_url)

**Lemma 5.1.** $K$ is fibered.

**Proof.** Let $W = E(K) - \text{Int}N(Q)$. Then $\partial W$ consists of two copies of $Q$, $Q_0$ and $Q_1$ say, and an annulus $\delta$. We show that $W$ has a product structure $Q \times [0, 1]$ such that $Q \times \{0\} = Q_0$ and $Q \times \{1\} = Q_1$. Then the result immediately follows from this.
We see that $P \cap W$ consists of a 4-gon $D_4$ and $(m - 1) + (n - 1)$ bigons. See Figure 6. For such a bigon $D$, $\partial D \cap Q_0$ and $\partial D \cap Q_1$ correspond to edges of $G_Q$ and $\partial D \cap \delta$ is two spanning arcs of $\delta$. For example, if $D$ corresponds to the bigon face of $P$ between $a_1$ and $a_2$, then $\partial D \cap Q_0$ and $\partial D \cap Q_1$ correspond to $a_1$ and $a_2$, respectively. Cut $W$ along these bigons. Let $W'$ be the resulting manifold. That process cuts $Q_i$ into a disk for each $i$, since $Q_0$ is cut along arcs corresponding to $a_1, a_2, \ldots, a_{m-1}, b_1, b_2, \ldots, b_{n-1}$, and $Q_1$ is cut along $a_2, a_3, \ldots, a_m, b_2, b_3, \ldots, b_n$. By the irreducibility of $E(K)$, $W'$ is a 3-ball. Thus $W$ has a product structure as desired.

Thus $W$ is identified with $Q \times [0, 1]$, and $E(K)$ is identified with a mapping torus $Q \times [0, 1]/(x, 1) = (f(x), 0)$, where $f$ is a homeomorphism of $Q$. Let $Q_i$ denote $Q \times \{i\}$ in $W = Q \times [0, 1]$. In fact, it is convenient to regard $f$ as the map from $Q_1$ to $Q_0$.

Let us keep using the notation in the proof of Lemma 5.1. To clarify the argument, we regulate $P$ in $E(K)$ up to isotopy. In the same way as [FH, Proposition 2.1], $P$ can be isotoped to be monotone except for just one saddle point in $\text{Int}P$ with respect to the bundle structure of $E(K)$. Furthermore, we may assume that $2m + 2n$ arcs on $\delta$ coming from $\partial P$ and $m + n - 2$ disks in $W = Q \times [0, 1]$ coming from the bigon faces of $G_P$ are all saturated (that is, the unions of $I$-fibers) with respect to the product structure of $W$. Finally, we isotope the 4-gon $D_4$ of $P \cap W$ such that $\pi|_{D_4}$ is an embedding except for four arcs on $\delta$, where $\pi : W = Q \times [0, 1] \to Q_1$ denotes the natural projection.

Hereafter, we regard the edges $a_i, b_j$ of $G_Q$ as the arcs on $Q_0$. This means that each $a_i, 1 \leq i \leq m - 1$, appears as the intersection of $Q_0$ and the disk which corresponds to a bigon face of $G_P$ between $a_i$ and $a_{i+1}$, and the arc $a_m$ is one of the arcs of $D_4 \cap Q_0$. Further, we set $a'_i = \pi(a_i)$ on $Q_1$. Then $a_{i+1} = f(a'_i)$ holds for each $i = 1, 2, \ldots, m - 1$.

**Lemma 5.2.** $K$ is composite.

**Proof.** Let us introduce two more arcs on $Q_0$ as follows.

First, let $a_{m+1} = f(a'_m)$. Recall that the endpoints of $a_i$ are labeled by $i$ and $m + i$ and those of $b_j$ are labeled by $2m + j$ and $2m + n + j$. Thus the action of $f$ is cyclic on the set of the endpoints of $a_i, b_j$, and so the labels of the endpoints of $a_{m+1}$ are $m + 1$ and $2m + 1$.

**Claim 5.3.** $a_{m+1}$ is disjoint from $a_2, a_3, \ldots, a_m$ and meets $a_1$ in only the endpoint with label $m + 1$.

**Proof.** Clearly, $a_m$ is disjoint from $a_i$, and so $a'_m$ is disjoint from $a'_i, 1 \leq i \leq m - 1$. This implies that $a_{m+1} = f(a'_m)$ is disjoint from $a_{i+1} = f(a'_i)$ for $1 \leq i \leq m - 1$. Furthermore, since $D_4 \cap Q_0 = a_m \cup b_n$, $D_4 \cap Q_1 = f^{-1}(a_1) \cup f^{-1}(b_1)$ and $\pi|_{D_4}$ is embedding except for four arcs on $\delta$, $a'_m$ meets $f^{-1}(a_1)$
which is obtained by shrinking corresponding to those on to be the homeomorphism induced from $f$ from $Q$ to a single point. Hence $a_{m+1} = f(a_m')$ meets $a_1$ in only the endpoint with label $m + 1$. \hfill \Box$

Next, we give an orientation to each edge of $G_Q$ so that it runs from the endpoint with smaller label to the other, and let $e$ be the arc on $Q_0$ obtained as the product $a_1 * a_{m+1}$. By the observations above, the endpoints of $e$ have the labels of 1 and $2m + 1$, and $e$ is disjoint from $a_2, \ldots, a_m$.

**Claim 5.4.** $e$ is separating and essential in $Q_0$.

**Proof.** Recall that $G_Q$ has just three disk faces. One of the disk faces, denoted by $D_a$, is bounded by the edges $a_1, a_2, \ldots, a_m$ together with subarcs of $\partial Q_0$. Another one $D_b$ is bounded by the edges $b_1, b_2, \ldots, b_n$ and the other one $D_{ab}$ is bounded by the edges $a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_n$ together with subarcs of $\partial Q_0$.

Since the labels of the endpoints of $a_{m+1}$ are $m + 1$ and $2m + 1$, the arc $a_{m+1}$ is contained in the disk face $D_{ab}$. Thus $e$ is also contained in $D_{ab}$, and so it is a diagonal arc which separates $D_{ab}$. Moreover, since the labels of the endpoints of $e$ are 1 and $2m + 1$, the endpoints separate $\partial D_{ab}$ into two parts one of which contains $a_i$’s only and the other contains $b_j$’s only. This indicates that $e$ is separating and essential on $Q_0$. \hfill \Box

Now, we consider the closed surface $\overline{Q}$ which is obtained by shrinking $\partial Q_i$ to a single point $y_i$ for $i = 0, 1$. We abuse the notations for the arcs and the faces on $\overline{Q}$; corresponding to those on $Q_i$. Let $\overline{f}$ be the homeomorphism from $\overline{Q}_1$ to $\overline{Q}_0$ induced from $f$.

**Claim 5.5.** $\overline{f}(e')$ is isotopic to $e$ fixing $y_0$, where $e' = \pi(e)$ in $Q_1$.

**Proof.** Let $[a_i], [a_i']$, $1 \leq i \leq m + 1$, and $[e], [e']$ be the elements of $\pi_1(\overline{Q}_0, y_0)$ and $\pi_1(\overline{Q}_1, y_1)$ represented by the corresponding ones. Let $R$ be the polygon bounded by $e$ and $a_1, a_2, \ldots, a_m$ in $D_{ab}$. Then, under the above setting, $\partial R$ is represented as

$$a_1 * a_2^{-1} * a_3 * a_4^{-1} * \cdots * a_m * e^{-1}.$$  

Then, this gives the relation

$$[e] = [a_1][a_2]^{-1}[a_3][a_4]^{-1}\cdots [a_m],$$

and so we have

$$[e'] = [a_1'][a_2']^{-1}[a_3'][a_4']^{-1}\cdots [a_m'].$$

Also $\partial D_a$ is represented as

$$a_1 * a_m * a_{m-1}^{-1} * a_{m-2} * a_{m-3}^{-1} * \cdots * a_2^{-1}.$$  

This gives

$$[a_1] = [a_2][a_3]^{-1}[a_4][a_5]^{-1}\cdots [a_m].$$
Let $f_*: \pi_1(\overline{Q}_1, y_1) \to \pi_1(\overline{Q}_0, y_0)$ be the homomorphism induced from $f$. Then

$$f_*(\langle e' \rangle) = f_*([a'_1][a'_2][a'_3][a'_4]^{-1} \ldots [a'_m])$$

$$= [a_2][a_3][a_4][a_5]^{-1} \ldots [a_m][a_{m+1}]$$

$$= [a_1][a_{m+1}] = [e].$$

Therefore two loops $e$ and $f(e')$ are homotopic on $\overline{Q}_0$ fixing $y_0$, and hence isotopic. □

As a result, we can obtain an essential, separating annulus in $E(K)$, each of whose boundary circles meets the longitude of $K$ in a single point, from $e \times [0, 1] \subset W$. By [BZ, Lemma 15.26], such an annulus comes from a decomposing sphere or a cabling annulus for $K$. This concludes that $K$ is composite. □

Proof of Theorem 1.2. By Lemma 5.2, $K$ is composite. Then by [T2], $K$ is the connected sum of two 2-cabled knots $K_1$ and $K_2$. Let $K_i$ be the $(2, m_i)$-cable of a knot $\tilde{K}_i$ for $i = 1, 2$. Then $r = 2m_1 + 2m_2$ [T2]. Also,

$$g(K) = \frac{|m_1| - 1}{2} + \frac{|m_2| - 1}{2} + 2g(\tilde{K}_1) + 2g(\tilde{K}_2)$$

by [S]. Since $|r| = 4g(K) + 4$,

$$2|m_1 + m_2| = 2|m_1| + 2|m_2| + 8g(\tilde{K}_1) + 8g(\tilde{K}_2).$$

Thus $m_1$ and $m_2$ have the same sign and $g(\tilde{K}_1) = g(\tilde{K}_2) = 0$, and hence $K_i$ is the $(2, m_i)$-torus knot. This completes the Proof of Theorem 1.2. □

Finally, we give the examples of hyperbolic knots which show that the estimation of Corollary 1.3 is sharp for each genus $g$.

\[ \text{Figure 7.} \]

Example 5.6. For genus one case, the figure eight knot is such an example as mentioned in Section 1. Let $n \geq 2$ and let $K$ be the $(2, 3, 2n - 3)$-pretzel knot. (When $n = 2$, $K$ is 2-bridge, and it is $6_2$ in the knot table [R].) Then
$K$ is hyperbolic [Kw], and it obviously bounds a once-punctured Klein bottle whose boundary has the slope $4n$. Also, $K$ has genus $n$, since the Seifert surface shown in Figure 7 has minimal genus by [G1, G2].

References


[H] J. Hoffman, There are no strict great $x$-cycles after a reducing or $P^2$ surgery on a knot, J. Knot Theory Ramifications, 7 (1998), 549-569, MR 99h:57010a, Zbl 0912.57009.


Received October 22, 2001 and revised July 20, 2002. The first author is supported in part by JSPS Research Fellowships for Young Scientists, and the second author is partially supported by Japan Society for the Promotion of Science, Grant-in-Aid for Encouragement of Young Scientists 12740041, 2001.

**DEPARTMENT OF INFORMATION AND COMPUTER SCIENCES**
**DEPARTMENT OF MATHEMATICS AND MATHEMATICS EDUCATION**

**FACULTY OF SCIENCE, NARA WOMEN’S UNIVERSITY**
**HIROSHIMA UNIVERSITY**

**KITA-UOYA NISHIMACHI**
**1-1-1 KAGAMIYAMA**

**NARA 630-8506**
**HIGASHI-HIROSHIMA 739-8524**

**JAPAN**
**JAPAN**

**E-mail address:** ichihara@vivaldi.ics.nara-wu.ac.jp

**E-mail address:** teragai@hiroshima-u.ac.jp
MULTI-VARIABLE POLYNOMIAL SOLUTIONS TO PELL’S EQUATION AND FUNDAMENTAL UNITS IN REAL QUADRATIC FIELDS

J. Mc Laughlin

Solving Pell’s equation is of relevance in finding fundamental units in real quadratic fields and for this reason polynomial solutions are of interest in that they can supply the fundamental units in infinite families of such fields.

In this paper an algorithm is described which allows one to construct, for each positive integer \( n \), a finite collection, \( \{ F_i \} \), of multi-variable polynomials (with integral coefficients), each satisfying a multi-variable polynomial Pell’s equation

\[
C_i^2 - F_i H_i^2 = (-1)^{n-1},
\]

where \( C_i \) and \( H_i \) are multi-variable polynomials with integral coefficients. Each positive integer whose square-root has a regular continued fraction expansion with period \( n + 1 \) lies in the range of one of these polynomials. Moreover, the continued fraction expansion of these polynomials is given explicitly as is the fundamental solution to the above multi-variable polynomial Pell’s equation.

Some implications for determining the fundamental unit in a wide class of real quadratic fields is considered.

1. Introduction.

Solving Pell’s equation is of relevance in finding fundamental units in real quadratic fields and for this reason polynomial solutions are interesting in that they can supply the fundamental units in infinite families of such fields.

There have been several papers written over the past thirty years which describe certain polynomials whose square roots have periodic continued fraction expansions which can be written down explicitly in terms of the coefficients and variables of the polynomials. See for example the papers of Bernstein [1], Levesque and Rhin [4], Madden [5], Van der Poorten [10] and Van der Poorten and Williams [11].

In this paper an algorithm is described which allows one to construct, for each positive integer \( n \), a finite collection of multi-variable Fermat-Pell polynomials which have all positive integers whose square-roots have a continued fraction expansion of period \( n + 1 \) in their range. If \( F_i := \)
The polynomials $F_i\left(t_0, t_1, \ldots, t_{\lfloor n+1 \rfloor}\right)$ are any one of these polynomials, the fundamental polynomial solution to the equation

\[ C_i^2 - F_i H_i^2 = (-1)^{n-1} \tag{1.1} \]

where $C_i$ and $H_i$ are polynomials in the variables $t_0, t_1, \ldots, t_{\lfloor n+1 \rfloor}$ can be found. Moreover, the continued fraction expansion of $\sqrt{F_i}$ can be written down when $t_1, \ldots, t_{\lfloor n+1 \rfloor} \geq 0$ and $t_0 > g_i\left(t_1, \ldots, t_{\lfloor n+1 \rfloor}\right)$, a certain rational function of these variables. Some implications for single-variable Fermat-Pell polynomials are discussed as are the implications for writing down the fundamental units in a wide class of real quadratic number fields.

**Definition.** A multi-variable polynomial $F := F(t_0, t_1, \ldots, t_k) \in \mathbb{Z}[t_0, t_1, \ldots, t_k]$, $k \geq 1$ is called a multi-variable Fermat-Pell polynomial if there exists polynomials $C := C(t_0, t_1, \ldots, t_k)$ and $H := H(t_0, t_1, \ldots, t_k) \in \mathbb{Z}[t_0, t_1, \ldots, t_k]$ such that either

\[ C_i^2 - F_i H_i^2 = 1, \quad \text{for all } t_i, \quad 0 \leq i \leq k, \text{ or} \]
\[ C_i^2 - F_i H_i^2 = -1, \quad \text{for all } t_i, \quad 0 \leq i \leq k. \tag{1.2} \]

Such a triple of polynomials $\{C, H, F\}$ satisfying Equation (1.2) constitute a multi-variable polynomial solution to Pell’s equation.

**Definition.** The multi-variable Fermat-Pell polynomial $F$ (as above) is said to have a multi-variable polynomial continued fraction expansion if there exists a positive integer $n$, a real constant $T$, a rational function $g(t_1, \ldots, t_k) \in \mathbb{Q}(t_1, \ldots, t_k)$ and polynomials $a_0 := a_0(t_0, t_1, \ldots, t_k) \in \mathbb{Z}[t_0, t_1, \ldots, t_k]$ and $a_j := a_j(t_1, \ldots, t_k) \in \mathbb{Z}[t_1, \ldots, t_k]$, $1 \leq j \leq n$, which take only positive integral values for integral $t_i \geq T, 1 \leq i \leq k$ and (possibly half-) integral $t_0 > g(t_1, \ldots, t_k)$ such that

\[ \sqrt{F} = [a_0; a_1, \ldots, a_n, 2a_0], \quad \text{for all } t_i \text{'s in the ranges stated, } 0 \leq i \leq k. \]

**Remarks.**

1. From the point of view of simplicity it would be desirable to replace the condition $t_0 \geq g(t_1, \ldots, t_k)$ by $t_0 \geq T$ but it will be seen that for the polynomials examined here that the former condition is more natural and indeed cannot be replaced by the latter condition.

These polynomials are called “Fermat-Pell polynomials” here to avoid confusion with “Pell Polynomials” and also because Fermat investigated the “Pell” equation.
The restriction that the $a_i(t_1, \ldots, t_k) \geq 0$, $1 \leq i \leq n$ may also seem artificial to some since negative terms can easily be removed from a continued fraction expansion (see, for example [10]) but this changes the period of the continued fraction so is avoided here.

It may also seem artificial to have $a_0$ depend on a variable $t_0$ while the other $a_i$’s do not but this will also be seen to occur naturally.

Finally, allowing $t_0$ to take half-integral values in some circumstances may also seem strange but this also will be seen to be natural and indeed necessary.

**Definition.** If, for all sets of integers $\{t'_0, t'_1, \ldots, t'_k\}$ satisfying $t'_0 \geq g(t'_1, \ldots, t'_k)$ and $t'_i \geq T$, $1 \leq i \leq k$,

$$X = C_i(t'_0, t'_1, \ldots, t'_k), \ Y = H_i(t'_0, t'_1, \ldots, t'_k)$$

constitutes the fundamental solution (in integers) to

$$X^2 - F_i(t'_0, t'_1, \ldots, t'_k)Y^2 = (-1)^{n-1}$$

then $(C_i(t_0, t_1, \ldots, t_k), H_i(t_0, t_1, \ldots, t_k))$ is termed the fundamental polynomial solution to Equation (1.1).

Standard notations are used:

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \ldots + \frac{1}{a_N}}} := a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \ldots + \frac{1}{a_N}}}.$$  

To save space this continued fraction is usually written $[a_0; a_1, \ldots, a_n]$. The infinite periodic continued fraction with initial non-periodic part $a_0$ and periodic part $a_1, \ldots, a_n, 2a_0$ is denoted by $[a_0; a_1, \ldots, a_n, 2a_0]$. The $i$-th approximant of the continued fraction $[a_0; a_1, \ldots, a_n]$ is denoted by $P_i/Q_i$.

Repeated use will be made of some basic facts about continued fractions, such as:

$$P_nQ_{n-1} - P_{n-1}Q_n = (-1)^{n-1},$$

$$P_{n+1} = a_{n+1}P_n + P_{n-1},$$

$$Q_{n+1} = a_{n+1}Q_n + Q_{n-1},$$

each of these relations being valid for $n = 1, 2, 3, \ldots$.

Before coming to the main problem, it is necessary to first solve a related problem on symmetric strings of positive integers.
2. A problem concerning symmetric sequences.

Question: For which symmetric sequences of positive integers \(a_1, \ldots, a_n\) do there exist positive integers \(a_0\) and \(D\) such that

\[
\sqrt{D} = [a_0; a_1, \ldots, a_n, 2a_0]?
\]

Let \(P_i/Q_i\) denote the \(i\)th approximant of the continued fraction

\[
0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \ldots + \frac{1}{a_n}}}}.
\]

By the well-known correspondence between convergents and matrices

\[
\begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}
\begin{pmatrix}
a_1 & 1 \\
1 & 0
\end{pmatrix}
\ldots
\begin{pmatrix}
a_n & 1 \\
1 & 0
\end{pmatrix}
= \begin{pmatrix}
P_n & P_{n-1} \\
Q_n & Q_{n-1}
\end{pmatrix}.
\]

Since the left side in the second equation is a symmetric sequence of symmetric matrices it follows that

\[
P_n = Q_{n-1}.
\]

Suppose \(\sqrt{D} = [a_0; a_1, \ldots, a_n, 2a_0] = a_0 + \beta\), where \(\beta = [0; a_1, \ldots, a_n, 2a_0]\) so that

\[
\beta = [0; a_1, \ldots, a_n, 2a_0 + \beta],
\]

\[
\Rightarrow \beta = \frac{(2a_0 + \beta)P_n + P_{n-1}}{(2a_0 + \beta)Q_n + Q_{n-1}} = \frac{\beta P_n + (2a_0 P_n + P_{n-1})}{\beta Q_n + (2a_0 Q_n + Q_{n-1})},
\]

\[
\Rightarrow \beta^2 Q_n + (2a_0 Q_n + Q_{n-1} - P_n)\beta - (2a_0 P_n + P_{n-1}) = 0,
\]

\[
\Rightarrow \beta^2 Q_n + (2a_0 Q_n)\beta - (2a_0 P_n + P_{n-1}) = 0, \text{ (by (2.3))}
\]

\[
\Rightarrow \sqrt{D} = a_0 + \beta = \sqrt{a_0^2 + \frac{2a_0 P_n + P_{n-1}}{Q_n}}.
\]

The problem now becomes one of determining for which symmetric sequences of positive integers \(a_1, \ldots, a_n\) does there exist positive integers \(a_0\) such that \((2a_0 P_n + P_{n-1})/Q_n\) is an integer.

**Theorem 1.** There exists a positive integer \(a_0\) such that \((2a_0 P_n + P_{n-1})/Q_n\) is an integer if and only if \(P_{n-1}Q_{n-1}\) is even.

**Proof.** \(\Leftarrow\) Suppose first of all that \(P_{n-1}Q_{n-1}\) is even. By Equation (1.3)

\[
P_n Q_{n-1} + (-1)^n = P_{n-1} Q_n.
\]
(i) Suppose \( n \) is even. Then \( P_n Q_{n-1} P_{n-1} + P_{n-1} = P_{n-1}^2 Q_n \). Choose \( t \) to be any integer or half-integer such that \( tQ_n \) is an integer and \( a_0 := Q_{n-1} P_{n-1} / 2 + tQ_n > 0 \). Then
\[
\frac{2a_0 P_n + P_{n-1}}{Q_n} = \frac{Q_{n-1} P_{n-1} P_n + 2tP_n Q_n + P_{n-1}}{Q_n} = 2tP_n + P_{n-1}^2.
\]

(ii) Similarly, in the case \( n \) is odd, \( -P_n Q_{n-1} P_{n-1} + P_{n-1} = -P_{n-1}^2 Q_n \). Choose \( t \) to be any integer or half-integer such that \( tQ_n \) is an integer and \( a_0 := -Q_{n-1} P_{n-1} / 2 + tQ_n > 0 \). In this case
\[
\frac{2a_0 P_n + P_{n-1}}{Q_n} = 2tP_n - P_{n-1}^2.
\]

\[\implies \text{Suppose next that } P_{n-1} \text{ and } Q_{n-1} \text{ are both odd and that there exists a positive integer } a_0 \text{ such that } (2a_0 P_n + P_{n-1})/Q_n \text{ is a positive integer, } m, \text{ say. Using (1.3) and (2.3) it follows that } Q_n \text{ is even. Then } 2a_0 P_n + P_{n-1} = mQ_n \implies P_{n-1} \text{ is even - a contradiction.} \]

**Remarks.**

(i) Note that this process gives all \( a_0 \) such that \( (2a_0 P_n + P_{n-1})/Q_n \) is an integer. Indeed,
\[
(2a_0 P_n + P_{n-1})/Q_n = k, \text{ an integer}
\]
\[\iff 2a_0 P_n Q_{n-1} = -P_{n-1} Q_{n-1} + kQ_n Q_{n-1},
\]
\[\iff 2a_0 (-1)^{n-1} = 2a_0 (P_n Q_{n-1} - P_{n-1} Q_n), \text{ (by (1.3))}
\]
\[\iff -P_{n-1} Q_{n-1} + Q_n (kQ_{n-1} - 2a_0 P_{n-1}),
\]
\[\iff a_0 = (-1)^{n-1} \left( \frac{-P_{n-1} Q_{n-1}}{2} + Q_n \frac{kQ_{n-1} - 2a_0 P_{n-1}}{2} \right).
\]

Notice also that if there is one such \( a_0 \) that there are infinitely many of them.

(ii) Notice that, with \( P_n, P_{n-1}, Q_n \) and \( Q_{n-1} \) as defined above, if there exists a positive integer \( D \) satisfying (2.1) then \( D = p(t_0) \), for some allowed \( t_0 \), where
\[
p(t) = \left( \frac{Q_{n-1} P_{n-1}}{2} + tQ_n \right)^2 + 2tP_n + P_{n-1}^2, \text{ (} n \text{ even),}
\]
\[
p(t) = \left( \frac{-Q_{n-1} P_{n-1}}{2} + tQ_n \right)^2 + 2tP_n - P_{n-1}^2, \text{ (} n \text{ odd).}
\]

The above theorem suggests a simple algorithm for deciding if, for a given symmetric sequence of positive integers \( a_1, \ldots, a_n \), there exist positive integers \( a_0 \) and \( D \) such that (2.1) holds. Notice that all that matters is the parity of the \( a_i \) so all calculations can be done in \( \mathbb{Z}_2 \). First of all define the
following matrices:

\[ J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad K = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \]

Convert the sequence \( a_1, a_2, \ldots, a_n \) to a sequence of \( J \)- and \( K \)-matrices, according to whether each \( a_i \) is odd (replace by a \( K \)) or even (replace by a \( J \)).

Prefix a \( J \)-matrix (to account for the initial 0 in the continued fraction (2.2)). Multiply this sequence together (modulo 2) using the facts that \( J^2 = K^3 = I \), and \( JK = K^2 J \).

The final matrix \( \equiv \begin{pmatrix} * & 1 \\ * & 1 \end{pmatrix} \mod 2 \iff \) there do not exist positive integers \( a_0 \) and \( D \) such that (2.1) holds.

**Example 1.** Do there exist positive integers \( a_0 \) and \( D \) such that

\[ \sqrt{D} = [a_0; 22, 34, 97, 32, 15, 17, 15, 32, 97, 34, 22, 2a_0]? \]

As described above convert the sequence 22, 34, 97, 32, 15, 17, 15, 32, 97, 34, 22 to a sequence of \( J \)- and \( K \)-matrices, prefix a \( J \)-matrix and multiply the sequence together:

\[
JJ \underbrace{JKJKJKK}_{KJK} JJ = JK(JK)JK = J K(K^2 J)JK = JK = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.
\]

Therefore there do exist positive integers \( a_0 \) and \( D \) such that

\[ \sqrt{D} = [a_0; 22, 34, 97, 32, 15, 17, 15, 32, 97, 34, 22, 2a_0]. \]

### 3. Multi-variable Fermat-Pell polynomials.

**Definition.** If \( \{a_1, \ldots, a_n\} \) is a symmetric zero-one sequence such that

\[
\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \prod_{i=1}^{n} \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} \not\equiv \begin{pmatrix} * & 1 \\ * & 1 \end{pmatrix} \mod 2
\]

then the sequence \( \{a_1, \ldots, a_n\} \) is termed a permissible sequence. Let \( r(n) \) denote the number of permissible sequences of length \( n \).

Note: It is not difficult to show that \( r(2m) = (-1)^m + 2^{m+1}/3 \) and that \( r(2m+1) = ((-1)^m + 5 \times 2^m)/3 \).

If \( D \) is a positive integer such that \( \sqrt{D} = [a_0; a_1, \ldots, a_n, 2a_0] \) then \( \{a_1, \ldots, a_n\} \mod 2 \) must equal one of the above permissible sequences and \( D \) is said to be associated with this permissible sequence. The collection of all positive integers associated with a particular permissible sequence is termed the parity class of this permissible sequence. Sometimes, if there is no danger of ambiguity, these collections of positive integers will be referred to simply as parity classes.
Theorem 2.

(i) For each positive integer \( n \) there exists a finite collection of multi-variable Fermat-Pell polynomials \( F_j\left(t_0, t_1, \ldots, t_{\left\lfloor \frac{n+1}{2} \right\rfloor}\right), 1 \leq j \leq r(n) \), such that each positive integer whose square root has a continued fraction expansion with period \( n + 1 \) lies in the range of exactly one of these polynomials. Moreover, these polynomials can be constructed.

(ii) These polynomials have a polynomial continued fraction expansion which can be explicitly determined.

(iii) The fundamental polynomial solution

\[
C = C_j \left(t_0, t_1, \ldots, t_{\left\lfloor \frac{n+1}{2} \right\rfloor}\right), \quad H = H_j \left(t_0, t_1, \ldots, t_{\left\lfloor \frac{n+1}{2} \right\rfloor}\right)
\]

exists and can be explicitly determined.

Proof. (i) The proof will be by construction.

Step 1. Find all permissible sequences. This will involve checking \( 2^{\left\lfloor \frac{n+1}{2} \right\rfloor} \) zero-one sequences in a way similar to Example 1 above.

Step 2. For each permissible sequence \( \{a_1, \ldots, a_n\} \) create a new symmetric polynomial sequence \( \{a_1(t_1), a_2(t_2), \ldots, a_{n-1}(t_2), a_n(t_1)\} \) by replacing each \( a_i \) and its partner \( a_{n+1-i} \) in the symmetric sequence by \( a_i(t_i) = a_{n+1-i}(t_i) = 2t_i + 1 \) if \( a_i = 1 \) and by \( a_i(t_i) = a_{n+1-i}(t_i) = 2t_i + 2 \) if \( a_i = 0 \). This new sequence will sometimes be referred to as the sequence \( \{a_1, \ldots, a_n\} \), if there is no danger of ambiguity. Each of the integer variables \( t_i \) (in the polynomial being constructed) will be allowed to vary independently over the range \( 0 \leq t_i < \infty \) and each of the new \( a_i \)'s will keep the same parity and stay positive.

Step 3. As in (2.2), form the continued fraction

\[
0 + \frac{1}{a_1(t_1)} + \frac{1}{a_2(t_2)} + \cdots + \frac{1}{a_{n-1}(t_2)} + \frac{1}{a_n(t_1)}
\]

and calculate \( P_n, Q_n, P_{n-1} \) and \( Q_{n-1} \) for this polynomial continued fraction, where these expressions are now polynomials in the \( t_i \)'s.

Step 4. Construct \( F_j := F_j(t_0, t_1, \ldots, t_{\left\lfloor \frac{n+1}{2} \right\rfloor}) \), the multi-variable Fermat-Pell polynomial corresponding to the particular parity sequence under consideration. This is simply done by defining

\[
F_j := \begin{cases} 
\left(\frac{Q_{n-1}P_{n-1}}{2} + t_0Q_n\right)^2 + 2t_0P_n + P_{n-1}^2, & \text{if } n \text{ even} \\
\left(-\frac{Q_{n-1}P_{n-1}}{2} + t_0Q_n\right)^2 + 2t_0P_n - P_{n-1}^2, & \text{if } n \text{ odd}
\end{cases}
\]

where \((-1)^{n+1}Q_{n-1}P_{n-1}/(2Q_n) < t_0 < \infty\) and \( t_0 \) can take half-integral values if \( Q_n \) is even and otherwise takes integral values.
Every positive integer whose square root has a continued fraction expansion with period $n + 1$ lies in the range of exactly one of these polynomials. That these polynomials are multi-variable Fermat-Pell polynomials follows from Equation (3.4) below.

(ii) With $t_0$ in the range given, then

$$\sqrt{F_j} = \left[ a_0 \left( t_0, t_1, \ldots, t_{\left\lfloor \frac{n+1}{2} \right\rfloor} \right); a_1(t_1), \ldots, a_n(t_1), 2a_0 \left( t_0, t_1, \ldots, t_{\left\lfloor \frac{n+1}{2} \right\rfloor} \right) \right],$$

for all $t_i \geq 0$. Here

$$a_0 = a_0 \left( t_0, t_1, \ldots, t_{\left\lfloor \frac{n+1}{2} \right\rfloor} \right) := \begin{cases} \frac{Q_{n-1}P_{n-1}}{2} + t_0 Q_n, & (n \text{ even}) \\ \frac{-Q_{n-1}P_{n-1}}{2} + t_0 Q_n, & (n \text{ odd}) \end{cases}.$$

(iii) Notice (using (1.3) and (2.3)) that

$$\left( a_0 Q_n + P_n \right)^2 - \left( a_0^2 + (2a_0 P_n + P_{n-1})/Q_n \right) Q_n^2 = (-1)^{n-1}.$$

To see that $(a_0 Q_n + P_n, Q_n)$ is the fundamental solution to (3.1), notice that

$$\sqrt{F_j} = \left[ a_0 \left( t_0, t_1, \ldots, t_{\left\lfloor \frac{n+1}{2} \right\rfloor} \right); a_1(t_1), \ldots, a_n(t_1), 2a_0 \left( t_0, t_1, \ldots, t_{\left\lfloor \frac{n+1}{2} \right\rfloor} \right) \right].$$

This has period $n + 1$ and the $n$th approximant is $a_0 + P_n/Q_n = (a_0 Q_n + P_n)/Q_n$ and by the theory of the Pell equation $(a_0 Q_n + P_n, Q_n)$ is the fundamental solution to (3.1).

As regards fundamental units in quadratic fields there is the following theorem on page 119 of [6]:

**Theorem 3.** Let $D$ be a square-free, positive rational integer and let $K = \mathbb{Q}(\sqrt{D})$. Denote by $\epsilon_0$ the fundamental unit of $K$ which exceeds unity, by $s$ the period of the continued fraction expansion for $\sqrt{D}$, and by $P/Q$ the $(s - 1)$-th approximant of it.

If $D \not\equiv 1 \mod 4$ or $D \equiv 1 \mod 8$, then

$$\epsilon_0 = P + Q\sqrt{D}.$$

However, if $D \equiv 5 \mod 8$, then

$$\epsilon_0 = P + Q\sqrt{D}$$

or

$$\epsilon_0^3 = P + Q\sqrt{D}.$$

Finally, the norm of $\epsilon_0$ is positive if the period $s$ is even and negative otherwise.
MULTI-VARIABLE POLYNOMIAL SOLUTIONS TO PELL'S EQUATION

It is easy, working modulo 4, to determine simple conditions (on $t_0$) which make $F_j \equiv 2$ or 3 mod 4 and thus to say further, for a particular set of choices of $t_1, \ldots, t_{\lfloor \frac{n+1}{2} \rfloor}$ and for all odd or even $t_0$, that if $F_j$ is square-free, then $a_0Q_n + P_n + \sqrt{F_j}Q_n$ is the fundamental unit in $\mathbb{Q}[\sqrt{F_j}]$. For example, suppose that $n$ is even and that the original $Q_{n-1}$ determined from the permissible zero-one sequence is also even (so that $P_{n-1}$ and $Q_n$ are both odd and $P_n = Q_{n-1}$ is even). Then the multi-variable form of $Q_{n-1}$ evaluated in Step 3 will also have all even coefficients. Suppose $Q_{n-1} \equiv c_0 + \sum t_i' \mod 2$. (Here $c_0$ may be 0 and the sum $\sum t_i'$ may contain some, all or none of the $t_i$'s.) It is easy to see that $F_j \left( t_0, t_1, \ldots, t_{\lfloor \frac{n+1}{2} \rfloor} \right) \equiv (c_0 + \sum t_i' + t_0)^2 + 1 \mod 4$. Even more simply, if the original $Q_{n-1}$ as in Step 1 is odd (here also the case $n$ is even is considered) then $P_{n-1}$ as evaluated in Step 3 is even and it is not difficult to show that in fact $P_{n-1} \equiv 2 \mod 4$ (since for $n$ even $P_nQ_{n-1} - P_{n-1}Q_n = -1$) and that $Q_n$ is odd, which leads to $F_j \left( t_0, t_1, \ldots, t_{\lfloor \frac{n+1}{2} \rfloor} \right) \equiv t_0^2 + 1 \mod 4$. Similar relations hold in the case where $n$ is odd.

The polynomials constructed in Theorem 2 take values in only one parity class, if all the variables are positive. However, given any two parity classes, there are multi-variable Fermat-Pell polynomials that take values in those two classes.

**Theorem 4.** Let $n$ be any fixed positive integer large enough so that the set of positive integers whose square roots have a continued fraction expansion of period $n + 1$ can be divided into more than one parity class.

(i) Given any two parity classes of integers whose square roots have continued fraction expansions of period $n + 1$, there are multi-variable Fermat-Pell polynomials, which can be constructed, that take values in both parity classes.

(ii) These polynomials have a polynomial continued fraction expansion which can be explicitly determined.

(iii) If $F = F \left( t_0, c, t_1, \ldots, t_{\lfloor \frac{n+1}{2} \rfloor} \right)$ is any such polynomial then the fundamental polynomial solution

$$C = C \left( t_0, c, t_1, \ldots, t_{\lfloor \frac{n+1}{2} \rfloor} \right), \quad H = H \left( t_0, c, t_1, \ldots, t_{\lfloor \frac{n+1}{2} \rfloor} \right)$$

to

$$C^2 - FH^2 = (\pm 1)^{n-1}$$

can be explicitly determined.

**Proof.** As in Step 2 in Theorem 2 a polynomial sequence $\{a_1, \ldots, a_n\}$ is created. Suppose $L_1 = \{b_1, \ldots, b_n\}$ and $L_2 = \{c_1, \ldots, c_n\}$ are the permissible sequences associated with the two parity classes. Let $i_1, \ldots, i_k$ be
those positions $\leq \lfloor \frac{n+1}{2} \rfloor$ at which the sequences agree. For each of these $i_r$’s set $a_{i_r}(t_{i_r}) = a_{n+1-i_r}(t_{i_r}) = 2t_{i_r} + 1$, if $c_{i_r}$ is odd and set $a_{i_r}(t_{i_r}) = a_{n+1-i_r}(t_{i_r}) = 2t_{i_r} + 2$, if $c_{i_r}$ is even. Subdivide the remaining positions (those positions $\leq \lfloor \frac{n+1}{2} \rfloor$ at which $L_1$ and $L_2$ differ) into two subsets: Those at which $L_1$ has a 0 and $L_2$ has a 1 and those at which $L_1$ has a 1 and $L_2$ has a 0.

Suppose $i_j$ is a position of the first kind. Let $a_{i_j}(c,t_{i_j}) = a_{n+1-i_j}(c,t_{i_j}) = c + 2 + 2t_{i_j}$. Repeat this for all the positions $i_j$ in this first set. Likewise, suppose $i_j$ is a position of the second kind. In this case let $a_{i_j}(c,t_{i_j}) = a_{n+1-i_j}(c,t_{\lfloor \frac{n+1}{2} \rfloor}) = c + 1 + 2t_{i_j}$. This is also repeated for all the positions $i_j$ in this second set. Step 3 and Step 4 are then carried out as above. The rest of the proof is identical to Theorem 2. Denote the polynomial produced by

$$F := F\left(t_0, c, t_1, \ldots, t_{\lfloor \frac{n+1}{2} \rfloor}\right).$$

As in Theorem 2, if $c$ and all the $t_i$’s are nonnegative, $1 \leq i \leq \lfloor \frac{n+1}{2} \rfloor$ and $t_0 > (-1)^{n+1}Q_n^{-1}P_{n-1}/(2Q_n)$ then

$$\sqrt{F} = [a_0; a_1, \ldots, a_n, 2a_0],$$

where the $a_i$’s, $1 \leq i \leq n$ are as defined just above and $a_0$ is as defined in Equation (3.3).

Under these conditions also the parity class of $F\left(t_0, c, t_1, \ldots, t_{\lfloor \frac{n+1}{2} \rfloor}\right)$ will depend only on the parity of $c$. As in Theorem 2 the fundamental polynomial solution to

$$C^2 - F\left(t_0, c, t_1, \ldots, t_{\lfloor \frac{n+1}{2} \rfloor}\right) H^2 = (-1)^{n-1}$$

is given by $C = a_0Q_n + P_n$, $H = Q_n$. □

4. A worked example.

As an example, consider those positive integers whose square-roots have continued fraction expansion with period of length 9. Thus the symmetric part of the period has length 8 and it is necessary to check the $2^4 = 16$ zero-one sequences to determine which are permissible. (This checking is done in essentially the same way as in Example 1 above.) There are 11 valid sequences:

0, 0, 0, 0, 0, 0, 0, 0
0, 0, 0, 0, 1, 0, 0, 0
0, 0, 1, 1, 1, 0, 0, 0
0, 1, 0, 0, 0, 0, 1, 0
The ninth of these is considered in more detail (each of the others can be dealt with in a similar way). For clarity the letters $a, b, c$ and $d$ are used instead of $t_1, t_2, t_3$ and $t_4$. Evaluating the continued fraction

\[
0 + \frac{1}{2a + 1 + \frac{1}{2b + 2 + \frac{1}{2c + 1 + \frac{1}{2d + 1 + \frac{1}{2c + 1 + \frac{1}{2b + 2 + \frac{1}{2a + 1}}}}}}}
\]

it is found that

\[
P_8 = Q_7 = -1 - 2d + 2 (3 + 4a + 2b + 4ab)(4 + 3b + 4c + 4bc + 6d + 4bd + 8cd + 8bcd) + 4 (3 + 2b + 4(1 + b)c)(2b + 3c + 2bc + a(3 + 2b + 4(1 + b)c)) \times (1 + 2d + 2d^2),
\]

\[
P_7 = 4(1 + b)(4 + 3b + 4c + 4bc + 6d + 4bd + 8cd + 8bcd) + 2 (3 + 2b + 4(1 + b)c)^2 (1 + 2d + 2d^2)\]

and

\[
Q_8 = 8 (2 + b + 3c + 2bc + a(3 + 2b + 4(1 + b)c))^2 (1 + 2d + 2d^2) + (3 + 4a + 2b + 4ab)(3 + 4c + 4d + 8cd + 2 + 4a)(4 + 3b + 4c + 4bc + 6d + 4bd + 8cd + 8bcd).\]

Since $n$ is 8 (even) and $Q_8$ is odd (so $t_0$ cannot take half-integer values), in this case $F_9(t_0, a, b, c, d)$ is defined by

\[
F_9(t_0, a, b, c, d) = (Q_7P_7/2 + t_0Q_8)^2 + 2t_0P_8 + P_7^2
\]
and
\[ \sqrt{F_9(t_0, a, b, c, d)} = (Q_7 P_7/2 + t_0 Q_8; 2a + 1, 2b + 2, 2c + 1, 2d + 1, 2d + 1), \]
this expansion being valid for all \( a, b, c, d \geq 0 \) and all \( t_0 > -Q_7 P_7/(2Q_8) \) and in particular for all \( t_0 \geq 0 \). In these ranges
\[ C = (Q_7 P_7/2 + t_0 Q_8) Q_8 + P_8, \quad H = Q_8 \]
gives the fundamental polynomial solution to
\[ C^2 - F_9 H^2 = -1. \]
\[ F_9(t_0, a, b, c, d) = (Q_7 P_7/2 + t_0 Q_8)^2 + 2t_0 P_8 + P_7^2 \equiv (1 + t_0^2) \mod 4, \]
so that if \((Q_7 P_7/2 + t_0 Q_8)^2 + 2t_0 P_8 + P_7^2 \) is a square-free number for some particular
\( a, b, c, d \geq 0 \) and some odd \( t_0 > -Q_7 P_7/(2Q_8) \), then
\[ (Q_7 P_7/2 + t_0 Q_8) Q_8 + P_8 + \sqrt{(Q_7 P_7/2 + t_0 Q_8)^2 + 2t_0 P_8 + P_7^2 Q_8} \]
is the fundamental unit in \( \mathbb{Q} \left( \sqrt{(Q_7 P_7/2 + t_0 Q_8)^2 + 2t_0 P_8 + P_7^2} \right) \).

5. Mystification, Fermat-Pell polynomials of a single variable and more on odd-even.

Clearly it is possible to “mystify” this process by replacing each \( t_i \) by some polynomial \( g_i(t_i) \) taking only positive values or by replacing \( 2t_i \) (recalling that the continued fraction expansion contains only terms like \( 2t_i + 1 \) or \( 2t_i + 2 \)) by some polynomial \( g_i(t_i) \) taking only even nonnegative values or by setting \( t_i = t_i(X_1, X_2, \ldots, X_k), 1 \leq i \leq \left\lfloor \frac{n+1}{2} \right\rfloor \), a polynomial in the \( X_j \)'s taking only positive values, where the \( X_j \)'s can be independent variables and \( k \) can be as large as desired and so on.

Finally of course one can obtain single-variable Fermat-Pell polynomials by replacing the original variables \( t_0, t_i, 1 \leq i \leq \left\lfloor \frac{n+1}{2} \right\rfloor \) by polynomials in a single variable. If it is desired that the period of the continued fraction expansion of the new single-variable Fermat-Pell polynomial should stay the same as that of the originating multi-variable polynomial then the domain of the single variable should be restricted so that the polynomials replacing each of the \( t_i \)'s take only positive values as in the multi-variable case and the polynomial replacing \( t_0 \) must be such that the \( a_0 \) term stays positive for all allowed values of the new single variable.

For example, letting \( a = s, b = 0, c = s, d = 0 \) and \( t_0 = s \) in the polynomial (4.2) above produces the single-variable Fermat-Pell polynomial
\[ g(s) = 639557 + 6858268 s + 33078145 s^2 + 94534688 s^3 + 177380352 s^4 + 228442240 s^5 + 204593408 s^6 + \]
which has the continued fraction expansion (valid for all \( s \geq 0 \))

\[
\sqrt{g(s)} = [799 + 4289 s + 9184 s^2 + 9856 s^3 + 5312 s^4 + 1152 s^5; \\
2s + 1, 2, 2s + 1, 1, 1, 2s + 1, 2, 2s + 1, \\
2(799 + 4289 s + 9184 s^2 + 9856 s^3 + 5312 s^4 + 1152 s^5)].
\]

\( g(s) \equiv (1 + s^2) \mod 4 \) so when \( s \) is odd and positive and \( g(s) \) is square-free

\[
51982 + 534625 s + 2429840 s^2 + 6408000 s^3 + \\
10812928 s^4 + 12115200 s^5 + 9019392 s^6 + 4304896 s^7 + 1196032 s^8 + \\
147456 s^9 + \sqrt{g(s)}(65 + 320 s + 576 s^2 + 448 s^3 + 128 s^4)
\]

is the fundamental unit in \( \mathbb{Q}[\sqrt{g(s)}] \). For example, letting \( s = 1 \) gives that

\[
47020351 + 1537 \sqrt{935888258}
\]

is the fundamental unit in \( \mathbb{Q}[\sqrt{935888258}] \).

Starting with the continued fraction

\[
0 + \frac{1}{2a + 1 + \frac{1}{2b + 2 + \frac{1}{c + 2d + 1 + \frac{1}{2d + 1 + \frac{1}{2d + 1 + \frac{1}{c + 2c + \frac{1}{2b + 2 + \frac{1}{2a + 1}}}}}}}}
\]

and following the same steps as above with the continued fraction (4.1) a multi-variable Fermat-Pell polynomial is developed which takes values in the parity classes associated with permissible sequences 7 and 9. Letting \( a = b = d = e = t = 0 \) one gets the single-variable Fermat-Pell polynomial

\[
g(c) = 4325 + 28140 c + 83652 c^2 + 147440 c^3 + 168000 c^4 + \\
126528 c^5 + 61504 c^6 + 17664 c^7 + 2304 c^8
\]

with continued fraction expansion

\[
\sqrt{g(c)} = [65 + 214c + 288c^2 + 184c^3 + 48c^4; \\
1, 2, c, 1, 1, c, 2, 1, 2(65 + 214c + 288c^2 + 184c^3 + 48c^4)],
\]

valid for \( c \geq 1 \).

Every Fermat-Pell polynomial in one variable, \( s \) say, that eventually has a continued fraction expansion of fixed period length can be found from (3.2), if it takes values in only one parity class for all sufficiently large \( s \), and from (3.6), if it takes values in two parity class for all sufficiently large \( s \). (Recall Remark (i) after Theorem 1.)

Of course none of this does anything to answer Schinzel’s question of whether every Fermat-Pell polynomial in one variable has a continued fraction expansion. Neither does it provide a criterion (such as Schinzel’s in the degree-two case) for deciding if a polynomial of arbitrarily high even degree is a Fermat-Pell polynomial. Perhaps it raises another question - Does every multi-variable Fermat-Pell polynomial have a continued fraction expansion? Does every multi-variable Fermat-Pell polynomial have a continued fraction expansion, assuming every Fermat-Pell polynomial in one variable does?

References


Received April 13, 2001 and revised July 16, 2002.

Mathematics Department
Trinity College
300 Summit Street
Hartford, CT 06106-3100
E-mail address: james.mclaughlin@trincoll.edu
THE FUNDAMENTAL GROUP OF THE COMPLEMENT OF A RESULTANT HYPERSURFACE

ICHIRO SHIMADA

Dedicated to Professor Tatsuo Suwa on his sixtieth birthday

We prove that the complement of a generalized resultant hypersurface has an abelian fundamental group.

1. Introduction

Let $X$ be a non-singular irreducible complex projective variety of dimension $n \geq 1$, and let $L_0, \ldots, L_n$ be very ample line bundles on $X$. We denote by $V_\nu$ the vector space $H^0(X, L_\nu)$, and set

$$V := V_0 \times \cdots \times V_n.$$ 

For $f_\nu \in V_\nu$, we put

$$(f_\nu) := \{ x \in X \mid f_\nu(x) = 0 \}.$$ 

The resultant variety $R$ of $V$ is defined to be

$$\{ f = (f_0, \ldots, f_n) \in V \mid (f_0) \cap \cdots \cap (f_n) \neq \emptyset \}.$$ 

It is known that $R$ is an irreducible hypersurface of $V$ ([GKZ, Chapter 3, Proposition 3.1]). Therefore we will call $R$ the resultant hypersurface.

When $X$ is the $n$-dimensional projective space $\mathbb{P}^n$, the resultant hypersurface $R$ is the classical resultant of $(n+1)$ forms in $(n+1)$-variables. See [GKZ] or [CLO] for other properties of the resultant hypersurfaces.

In this paper, we prove the following:

**Theorem 1.** The fundamental group of $V \setminus R$ is an infinite cyclic group.

In the case where $X = \mathbb{P}^1$, Theorem 1 follows from the result of [C], in which Choudary showed that the classical resultant hypersurface $R_{p,q}$ of polynomials of degree $p$ and $q$ has only normal crossings as its singularities in codimension 1, and proved the commutativity of $\pi_1(\mathbb{C}^{p+q} \setminus R_{p,q})$ by Zariski hyperplane section theorem [Z] and Fulton-Deligne’s Theorem ([D], [F], [FL]) on Zariski conjecture.

The generalized resultant hypersurface $R$ can have singularities in codimension 1 worse than normal crossings. For example, let $X \subset \mathbb{P}^2$ be a non-singular projective plane curve of degree $d \geq 3$, and let $L_0$ and $L_1$ be
the line bundles corresponding to a hyperplane section of $X$ in $\mathbb{P}^2$. Then a general fiber of the projection $R \to V_0$ consists of $d$ hyperplanes in $V_1$ passing through a fixed linear subspace of codimension 2.

In fact, as the proof in the next section shows, the case where we cannot apply Fulton-Deligne’s Theorem in a straightforward way (combined with Nori’s lemma [N, Lemma 1.5 (C)] and Zariski hyperplane section theorem) is always reduced to this example.

The fundamental group of the complement to the discriminant hypersurface of a linear system $|L|$ on a non-singular complex projective variety $X$ was studied by Dolgachev and Libgober in [DL]. We will explain the relation between the resultant hypersurface and the discriminant hypersurface in the case where $X = \mathbb{P}^n$ and $L = \mathcal{O}_X(d)$, where $n \geq 2$ and $d \geq 2$. We put $L_0 := L$ and $L_i := \mathcal{O}_X(d - 1)$ ($i = 1, \ldots, n$). The discriminant hypersurface $D \subset |L_0|$ is the projectivization of the hypersurface

$$\tilde{D} := \{f_0 \in V_0 \mid f_0 = 0 \text{ or } (f_0 \neq 0 \text{ and the divisor } (f_0) \text{ is singular})\}$$

in the vector space $V_0$ of homogeneous polynomials of degree $d$ in $(n + 1)$-variables. Let $(x_0 : x_1 : \cdots : x_n)$ be a homogeneous coordinate system of $X = \mathbb{P}^n$. We define a linear map $\varphi$ from $V_0$ to $V$ by

$$\varphi(f_0) := \left(f_0, \frac{\partial f_0}{\partial x_1}, \ldots, \frac{\partial f_0}{\partial x_n}\right).$$

Then we have

$$\tilde{D} = \varphi^{-1}(\varphi(V_0) \cap R);$$

that is, the discriminant hypersurface $\tilde{D}$ is a linear section of the resultant hypersurface $R$. Note that, since the image $\varphi(V_0)$ of $\varphi$ is not a general linear subspace of $V$, the non-commutativity of $\pi_1(|L_0| \setminus D)$ for many $n$ and $d$ (for example, see [DL, Section 4]) does not contradict to our theorem.

The author would like to thank the referee for many helpful comments on the first version of this paper.

2. Proof of Theorem 1.

First note that it is enough to prove that $\pi_1(V \setminus R)$ is abelian, because $R$ is irreducible.

For $\nu$ with $0 \leq \nu \leq n$, we put

$$V'_\nu := V_0 \times \cdots \times V_\nu, \quad V''_\nu := V_{\nu+1} \times \cdots \times V_n,$$

and denote by

$$\bar{p}_\nu : V \to V''_\nu$$

the natural projection. For a point $g$ of $V''_\nu$, we denote by $R_\nu(g)$ the intersection of $R$ with the fiber $\bar{p}_\nu^{-1}(g)$, and consider $R_\nu(g)$ as a Zariski closed
subset of $V'_\nu$. When $\nu = n$, $V''_n$ is the zero-dimensional vector space $\{0\}$, and we have $R_n(0) = R$. Let
\[ p_{\nu} : V \setminus R \to V''_\nu \]
be the restriction of $\bar{p}_{\nu}$ to $V \setminus R$. Then we have
\[ p_{\nu}^{-1}(g) = V'_\nu \setminus R_{\nu}(g). \]

Claim 2. If $g \in V''_\nu$ is general, the inclusion of $p_{\nu}^{-1}(g)$ into $V \setminus R$ induces a surjective homomorphism from $\pi_1(V'_\nu \setminus R_{\nu}(g))$ to $\pi_1(V \setminus R)$.

Proof of Claim 2. For $g = (g_{\nu+1}, \ldots, g_n) \in V''_\nu$, let $W_{\nu}(g)$ denote the subscheme of $X$ defined by
\[ g_{\nu+1} = \cdots = g_n = 0, \]
which is of dimension $\nu$ if $g$ is general in $V''_\nu$. We consider the universal family
\[ W_{\nu} \xrightarrow{\psi_{\nu}} X \]
\[ \phi_{\nu} \]
\[ V''_\nu \]
of the subschemes $W_{\nu}(g)$, where
\[ W_{\nu} := \{(g, x) \in V''_\nu \times X \mid g_{\nu+1}(x) = \cdots = g_n(x) = 0\}. \]
The projection $\psi_{\nu} : W_{\nu} \to X$ is smooth, and every fiber of $\psi_{\nu}$ is a linear subspace of $V''_\nu$ with codimension $n - \nu$. Hence $W_{\nu}$ is non-singular, irreducible and of dimension equal to $\dim V''_\nu + \nu$. On the other hand, the projection $\phi_{\nu} : W_{\nu} \to V''_\nu$ is surjective. Therefore there exists a Zariski closed subset $\Xi$ of $V''_\nu$ with codimension $\geq 2$ such that
\[ \dim W_{\nu}(g) = \nu \quad \text{for all} \quad g \in V''_\nu \setminus \Xi. \]
If $g \in V''_\nu \setminus \Xi$, then $R_{\nu}(g)$ is a proper Zariski closed subset of $V''_\nu$.

A general fiber of $p_{\nu} : V \setminus R \to V''_\nu$ is irreducible. If $g \in V''_\nu \setminus \Xi$, then $p_{\nu}^{-1}(g)$ has at least one point at which $p_{\nu}$ is smooth. Therefore Claim 2 follows from Nori’s lemma [N, Lemma 1.5 (C)].

We choose and fix a general point
\[ g = (g_1, \ldots, g_n) \]
of $V''_0$. We put
\[ d := c_1(L_1)c_1(L_2)\cdots c_1(L_n), \]
\[ d' := c_1(L_0)c_1(L_2)\cdots c_1(L_n), \]
where $c_1$ denote the first Chern class. Both of $d$ and $d'$ are positive integers. Then $W_0(g)$ consists of $d$ distinct points $a_1, \ldots, a_d$ of $X$, and $R_0(g)$ consists of $d$ distinct hyperplanes $H_1, \ldots, H_d$ of $V_0 = V_0$, where
\[ H_i := \{f_0 \in V_0 \mid f_0(a_i) = 0\}. \]
If \( d \leq 2 \), then \( \pi_1(V_0 \setminus R_0(g)) \) is obviously abelian. Hence \( \pi_1(V \setminus R) \) is abelian by Claim 2. Suppose that \( \dim V_\nu = 2 \) for some \( \nu \). Then we have \( n = 1 \), \( X = \mathbb{P}^1 \) and \( \deg(L_0) = 1 \). Interchanging \( L_0 \) and \( L_1 \), we will have \( d = 1 \), and can show the commutativity of \( \pi_1(V \setminus R) \) by the above argument. From now on, we will assume

\[
\dim V_\nu \geq 3 \quad \text{for} \quad \nu = 0, \ldots, n.
\]

Moreover, by interchanging \( L_0 \) and \( L_1 \) if necessary, we can assume

\[
d' \leq d.
\]

By the above argument, we can also assume

\[
3 \leq d.
\]

Suppose that \( R_0(g) \) satisfies the following:

\[
a_i \neq a_j \neq a_k \neq a_i \implies \dim(H_i \cap H_j \cap H_k) = \dim V_0 - 3.
\]

Let \( A \subset V_0 \) be a general affine plane. Then \( A \cap R_0(g) \) is a nodal affine plane curve consisting of \( d \) lines, no pairs of which are parallel. Hence \( \pi_1(A \setminus (A \cap R_0(g))) \) is abelian by Fulton-Deligne’s Theorem ([D], [F], [FL]) on Zariski conjecture. By Zariski hyperplane section theorem [Z], the inclusion

\[
A \setminus (A \cap R_0(g)) \hookrightarrow V_0 \setminus R_0(g)
\]

induces an isomorphism on the fundamental groups. Hence \( \pi_1(V_0 \setminus R_0(g)) \) is also abelian, and thus the commutativity of \( \pi_1(V \setminus R) \) follows from Claim 2.

Suppose, conversely, that there exist three distinct points \( a_i, a_j, a_k \) of \( W_0(g) \) such that

\[
\dim(H_i \cap H_j \cap H_k) = \dim V_0 - 2.
\]

Let \( U \) be a Zariski open dense subset of \( V_0'' \) containing the point \( g \) such that the projection \( \phi_0 : W_0 \rightarrow V_0'' \) is étale over \( U \). We have the monodromy action

\[
\mu : \pi_1(U, g) \rightarrow \mathfrak{S}(W_0(g))
\]

of \( \pi_1(U, g) \) on the finite set \( W_0(g) \), where \( \mathfrak{S}(W_0(g)) \) is the full symmetric group of \( W_0(g) \). Since the action \( \mu \) is doubly transitive, and the image of \( \mu \) contains a transposition, we see that \( \mu \) is surjective ([H, Uniform Position Lemma]). Since \( g \) is general in \( V_0'' \), we can conclude that (2.1) holds for any choice of distinct three points \( a_i, a_j, a_k \) of \( W_0(g) \). This means that, if a divisor \( D \in |L_0| \) of \( X \) contains distinct two points of \( W_0(g) \), then \( D \) contains every point of \( W_0(g) \).

When \( n = 1 \), we put \( h := 0 \in V_1'' = \{0\} \) and \( C := X \). In this case, we have \( p_1^{-1}(h) = V_1 \setminus R_1(h) = V \setminus R \). When \( n > 1 \), we put

\[
h := (g_2, \ldots, g_n),
\]
which is a general point of $V'^n_1$, and put

$$C := W_1(h).$$

We show that

$$\pi_1(p_1^{-1}(h)) = \pi_1(V'_1 \setminus R_1(h))$$

is abelian. The proof of Theorem 1 will then be completed by Claim 2.

First we will show that $C$ is a projective plane curve. The curve $C$ is non-singular and irreducible. The line bundles $L_0|C$ and $L_1|C$ on $C$ are very ample of degree $d'$ and $d$, respectively. Since the restriction $g_1|C$ of $g_1$ to $C$ is a general element of $H^0(C, L_1|C)$, and $d' \leq d$ has been assumed, we see from the above consideration that the following holds:

Let $D_1$ be a general divisor in the complete linear system $|L_1|C|$ on $C$. If a divisor $D_0$ in the complete linear system $|L_0|C|$ has at least two common points with $D_1$, then $D_0 = D_1$ holds.

In particular, we have $d = d'$ and $|L_1|C| = |L_0|C|$. We will denote by $P$ the dual projective space of the complete linear system $|L_1|C| = |L_0|C|$, and let

$$\Psi : C \to P$$

be the embedding of $C$ by $|L_1|C| = |L_0|C|$. Let $H$ be a general hyperplane of $P$. If $b_1$ and $b_2$ are points of $\Psi^{-1}(H)$, then $H$ is the only hyperplane containing $\Psi(b_1)$ and $\Psi(b_2)$. Therefore we have

$$\dim P = 2,$$

and $C$ can be regarded as a non-singular projective plane curve on $P$ via $\Psi$. The complete linear system $|L_1|C| = |L_0|C|$ is the linear system of intersections with lines in $P$.

We put

$$V_C := H^0(P, O_P(1)).$$

For $\lambda \in V_C$, let $(\lambda)$ denote the linear subspace of $P$ defined by $\lambda = 0$. We denote by $S$ the hypersurface

$$\{(\lambda_0, \lambda_1) \in V_C \times V_C \mid (\lambda_0) \cap (\lambda_1) \cap C \neq \emptyset\}$$

of $V_C \times V_C$, and put

$$(V_C \times V_C)^\circ := (V_C \times V_C) \setminus S.$$ 

The restriction map

$$(f_0, f_1) \mapsto (f_0|C, f_1|C)$$

gives a morphism

$$p_1^{-1}(h) = V'_1 \setminus R_1(h) \to (V_C \times V_C)^\circ,$$
which is locally trivial with fibers isomorphic to a vector space. Hence \( \pi_1(p_1^{-1}(h)) \) is isomorphic to \( \pi_1((V_C \times V_C)^0) \). Therefore it is enough to show the following:

**Claim 3.** The fundamental group of \( (V_C \times V_C)^0 \) is abelian.

**Proof of Claim 3.** We denote by

\[ \rho: (V_C \times V_C)^0 \to P \setminus C \]

the morphism given by

\[ \rho(\lambda_0, \lambda_1) := \text{the intersection point of the lines } (\lambda_0) \text{ and } (\lambda_1). \]

Then \( \rho \) is locally trivial, and its fiber is isomorphic to \( \text{GL}(2, \mathbb{C}) \). We choose a general line \( L_\infty \subset P \), and fix affine coordinates \((x, y)\) on \( P \setminus L_\infty \). Then \( \rho \) has a section \( \sigma: \]

\[ \sigma(a, b) := (x - a, y - b), \]

where \( x - a \) and \( y - b \) are considered as linear forms on \( P \). In particular, the fundamental group of \( (V_C \times V_C)^0 \setminus \rho^{-1}(L_\infty) \) is the semi-direct product

\[ \pi_1(\text{GL}(2, \mathbb{C})) \rtimes \pi_1(P \setminus (C \cup L_\infty)) \]

constructed from the monodromy action of \( \pi_1(P \setminus (C \cup L_\infty)) \) on \( \pi_1(\text{GL}(2, \mathbb{C})) \) associated with the section \( \sigma \). Since \( \pi_1(\text{GL}(2, \mathbb{C})) \cong \mathbb{Z} \) has a canonical positive generator, this monodromy action is trivial. Hence we have

\[ \pi_1((V_C \times V_C)^0 \setminus \rho^{-1}(L_\infty)) \cong \pi_1(\text{GL}(2, \mathbb{C})) \times \pi_1(P \setminus (C \cup L_\infty)). \]

Since \( C \cup L_\infty \) is a nodal curve, \( \pi_1(P \setminus (C \cup L_\infty)) \) is abelian. Therefore

\[ \pi_1((V_C \times V_C)^0 \setminus \rho^{-1}(L_\infty)) \]

is also abelian. Since the inclusion of \( (V_C \times V_C)^0 \setminus \rho^{-1}(L_\infty) \) into \( (V_C \times V_C)^0 \) induces a surjective homomorphism on the fundamental groups, we get the commutativity of \( \pi_1((V_C \times V_C)^0) \) \( \Box \)

**References**


RESULTANT HYPERSURFACE


Received September 2, 2001 and revised May 31, 2002.

DIVISION OF MATHEMATICS
GRADUATE SCHOOL OF SCIENCE
HOKKAIDO UNIVERSITY
SAPPORO 060-0810
JAPAN
E-mail address: shimada@math.sci.hokudai.ac.jp
ON THE ADDITIVITY OF THE THURSTON–BENNEQUIN INVARIANT OF LEGENDRIAN KNOTS

ICHIRO TORISU

In this article, we consider the maximal value of the Thurston–Bennequin invariant of Legendrian knots which topologically represent a fixed knot type in the standard contact 3-space and we prove a formula of the value under the connected sum operation of knots.

1. Introduction.

The standard contact structure $\xi_0$ on 3-space $\mathbb{R}^3 = \{(x, y, z)\}$ is the plane field on $\mathbb{R}^3$ given by the kernel of the 1-form $dz - ydx$. A Legendrian knot $K$ in the contact manifold $(\mathbb{R}^3, \xi_0)$ is a knot which is everywhere tangent to the contact structure $\xi_0$. The Thurston-Bennequin invariant $tb(K)$ of a Legendrian knot $K$ in $(\mathbb{R}^3, \xi_0)$ is the linking number of $K$ and a knot $K'$ which is obtained by moving $K$ slightly along the vector field $\frac{\partial}{\partial z}$. For a topological knot type $k$ in $\mathbb{R}^3$, the maximal Thurston-Bennequin invariant $mtb(k)$ is defined to be the maximal value of $tb(K)$, where $K$ is a Legendrian knot which topologically represents $k$. For any $k$, by the Bennequin’s inequality in [1], we know that $mtb(k)$ is an integer (i.e., not $\infty$). There are several computations of $mtb(k)$ (for example, see [3], [5], [8], [9], [10], [11]).

In this paper, we prove the following theorem:

**Theorem 1.1.** Let $k_1 \# k_2$ be the connected sum of topological knots $k_1$ and $k_2$ in $\mathbb{R}^3$. Then $mtb(k_1 \# k_2) = mtb(k_1) + mtb(k_2) + 1$.

**Remark 1.2.** After writing this paper, the author was informed that J. Etnyre and K. Honda [4] have also obtained a result on connected sum of Legendrian knots which extensively includes Theorem 1.1 and that T. Tanaka [12] have partially proved Theorem 1.1 by using a technique of algebraic knot theory.

2. Fronts.

Let $K$ be a Legendrian knot in $(\mathbb{R}^3, \xi_0 = \ker(dz - ydx))$. Then a diagram (i.e., projection) of $K$ in $xz$-plane is called front as in Figure 1.

A front does not have vertical tangents; generically, its only singularities are transverse double points and semicubical cusps. Note that the number
of the cusps is even. Since \( y = \frac{\partial z}{\partial x} \) along \( K \), the missing \( y \) coordinate is the slope of the front. Therefore the front of \( K \) is free from selftangencies, and, at a double point, the branch with a greater slope is higher along the \( y \) axis. Conversely such a diagram uniquely determines \( K \) as its front. So, as usual in knot theory, we identify a Legendrian knot \( K \) with its front, also denoted by \( K \).

The Thurston-Bennequin invariant \( tb(K) \) is computed in terms of the double points and cusps of its front. See Figure 2, where \( K \) is oriented and the choice of the orientation is irrelevant for the value of \( tb(K) \).

\[
tb = \# \begin{array}{c} \times \end{array} + \# \begin{array}{c} \times \end{array} - \# \begin{array}{c} \times \end{array} - \# \begin{array}{c} \times \end{array} -1/2 \# \text{ of cusps}
\]

For example, \( tb(K) = 1 \) for the front in Figure 1.

**Proposition 2.1.** For two topological knots \( k_1 \) and \( k_2 \), we have \( mtb(k_1 \# k_2) \geq mtb(k_1) + mtb(k_2) + 1 \).

**Proof.** Let \( K_1 \) and \( K_2 \) be Legendrian knots whose topological types are \( k_1 \) and \( k_2 \), respectively and \( mtb(k_1) = tb(K_1) \) and \( mtb(k_2) = tb(K_2) \). We also regard \( K_1 \) and \( K_2 \) as fronts. Further we can assume that \( K_1 \cap K_2 = \emptyset \) and \( K_1 \)}
Then we connect $K_1$ and $K_2$ by joining a right cusp of $K_1$ and a left cusp of $K_2$ as in Figure 4.

Figure 3.

Figure 4.
This procedure produces a Legendrian knot whose topological type is $k_1 \# k_2$ and Thurston-Bennequin invariant is $\text{mtb}(k_1) + \text{mtb}(k_2) + 1$. This completes the proof.  

3. Preliminaries from contact topology.

In this section, we recall some basic notions and theorems from recent 3-dimensional contact topology. In fact, the proof of Theorem 1.1 essentially relies on the previous foundational work of E. Giroux, Honda and Y. Eliashberg-M. Fraser. In particular, we assume the reader is familiar with convex surface theory started by Giroux in [6]. For details and proofs, see [2], [3], [6], [7], [8]. Let $\xi_n = \ker(\sin(2\pi nz)dx + \cos(2\pi nz)dy)$ be the contact structure on a solid torus $V = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 \leq \epsilon \}$, where $n \in \mathbb{N}$ and $\mathbb{R}^3$ is $\mathbb{R}^3$ modulo $z \mapsto z + 1$. The characteristic foliation on an embedded surface in a contact 3-manifold is the singular foliation defined by the intersection of the contact structure and the surface. The set of tangents of $\xi_n$ to $\partial V$ forms a disjoint union of two simple closed curves on $\partial V$, which are called Legendrian divides.

The next lemma is proved by a standard Darboux-type argument.

Lemma 3.1. For any Legendrian knot $K$ in $(\mathbb{R}^3, \xi_0)$, there exists a sufficiently small neighborhood $N(K)$ such that $(N(K), K, \xi_0)$ is isomorphic to $(V, \{(0, 0, z)\}, \xi_n)$ for some $n$.

As $\partial V$ is a convex surface (i.e., has a contact vector field transverse to $\partial V$), the following lemma can be proved by convex surface theory:

Lemma 3.2. Let $T$ be any embedded torus in $(\mathbb{R}^3, \xi_0)$ and $W$ a solid torus bounded by $T$. Suppose the characteristic foliation on $T$ is diffeomorphic to that on $\partial V$ and identifying these, the Legendrian divides on $T$ are isotopic to the core curve of $W$ through an isotopy in $W$. Then $(W, \xi_0)$ is isomorphic to $(V, \xi_n)$ for some $n$.

The following theorem on the classification of topologically trivial Legendrian knots due to Eliashberg-Fraser [2] is also needed for the proof of Theorem 1.1:

Theorem 3.3. Any topologically trivial Legendrian knot is Legendrian isotopic to one of standard forms expressed as fronts in Figure 5.

4. Proof of Theorem 1.1.

By Proposition 2.1, it is sufficient to show the converse inequality. Suppose $\hat{K}$ is a Legendrian knot in $(\mathbb{R}^3, \xi_0)$ whose topological type is the connected sum of $k_1$ and $k_2$ and its Thurston-Bennequin invariant is maximal. By Lemma 3.1, there exists a neighbourhood $N(\hat{K})$ of $\hat{K}$ such
that \((N(\hat{K}), \xi_0)\) is isomorphic to \((V, \xi_n)\) for some \(n\). Let \(B_1\) and \(B_2\) be 3-balls in \(\mathbb{R}^3\) such that \(B_1\) (resp. \(B_2\)) splits \(\hat{K}\) into the component corresponding to \(k_1\) (resp. \(k_2\)) and \(B_1 \cap B_2 = \emptyset\) (Figure 6).

Further, by convex surface theory, we can assume that (i) \(\partial B_1\) and \(\partial B_2\) are convex and (ii) \(\partial B_1 \cap \partial N(\hat{K})\) and \(\partial B_2 \cap \partial N(\hat{K})\) are Legendrian knots on \(\partial B_1\) and \(\partial B_2\), respectively and (iii) each dividing set on \(\partial B_i\) (i.e., the subset of \(\partial B_i\) consisting of tangents of \(\xi_0\) and a contact vector field defining the convex surface) intersects \(\partial B_i \cap N(\hat{K})\) as a diameter of the disk.
Then by Edge-Rounding Lemma due to Honda in [7], we have a solid torus $W$ such that (i) $W$ equals $B_1 \cup B_2 \cup N(\hat{K})$ except small neighbourhoods of $\partial B_1 \cap \partial N(\hat{K})$ and $\partial B_2 \cap \partial N(\hat{K})$ and (ii) $\partial W$ is a convex surface whose characteristic foliation is diffeomorphic to that of $\partial V$. By Lemma 3.2, it follows that $(W, \xi_0)$ is isomorphic to $(V, \xi_n)$ for some $n$. And notice that $W$ is unknotted in $\mathbb{R}^3$ and hence the core curve $K$ of $W$ which is Legendrian is also unknotted. Further, by a standard argument, we can assume that $K$ agrees with $\hat{K}$ in the region of $N(\hat{K}) - (B_1 \cup B_2)$. So by Theorem 3.3, $K$ is Legendrian isotopic to one of standard forms in Figure 5. Therefore $W$ is also identified with a small neighbourhood of that of the standard form. Further, by a homogeneous property of $V$ and a parallel translation of $W$, we can assume that a region of $W$ corresponding to $B_1$ (resp. $B_2$) lies in $\{(x,y,z)|x < 0\}$ (resp. $\{(x,y,z)|x > 0\}$). Then, identifying $\hat{K}$ with its front, we can divide $\hat{K}$ along a vertical line into Legendrian knots $K_1$ and $K_2$ corresponding to $k_1$ and $k_2$, respectively as the converse procedure in the proof of Proposition 2.1.

Counting the Thurston-Bennequin invariant of $K_1$ and $K_2$, we have $tb(\hat{K}) = mtb(k_1 \sharp k_2) = tb(K_1) + tb(K_2) + 1$. Therefore $mtb(k_1 \sharp k_2) \leq mtb(k_1) + mtb(k_2) + 1$.

This completes the proof of the main theorem.

References


Received July 12, 2001 and revised June 27, 2002.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
AKITA UNIVERSITY
1-1 TEUGATA GAKUEN-CHO
AKITA 010-8502
JAPAN

E-mail address: torisu@math.akita-u.ac.jp
REPRESENTATION OF TYPES AND 3-MANIFOLDS

Luis Gerardo Valdez Sánchez

According to theorems of C. Gordon, J. Luecke, and W. Parry, if a knot exterior $X$ has two distinct planar boundary slopes $r_1, r_2$, then at least one of the manifolds $X(r_1), X(r_2)$ has a connected summand $M$ with nontrivial torsion in first homology. The 3-manifolds $M$ obtained in this way, which we call t-manifolds, have special Heegaard splittings, or t-manifold structures. In this paper we study the topology of t-manifolds from the point of view of the homology presentation matrices induced by their t-manifold structures, classify all genus two t-manifold structures, and show that, under some conditions, one of the Dehn fillings of $X$ is a connected sum of t-manifolds and (at most) one prime non t-manifold summand.

1. Introduction.

Let $X$ be a knot exterior, i.e., a connected, compact, orientable, irreducible 3-manifold with torus boundary. For $r$ a slope in $\partial X$, let $X(r)$ be the manifold obtained by Dehn-filling $X$ along $r$. If $r$ and $s$ are two distinct planar boundary slopes on $\partial X$, then by [6, §4] at least one of the manifolds $X(r), X(s)$ contains a connected summand $M$ with nontrivial torsion in first homology. Such a manifold $M$ can be constructed as follows: Let $P, Q$ be essential planar surfaces in $X$ with boundary slope $r, s$, respectively; we may assume that $P, Q$ intersect transversely in planar graphs $G_P = P \cap Q \subset P$ and $G_Q = P \cap Q \subset Q$ without boundary parallel arcs, and that any circle component of $P \cap Q$ is essential in $P$ and $Q$. By [8] and [9], $\Delta(r, s) = 1$ and at least one of the two graphs, say $G_Q$, has a set of disk faces $\Sigma$ which represents all types. Let $\hat{P}$ be the surface obtained by capping each boundary component of $P$ with a disk, so that $\hat{P} = S^2$, and let $(E, \partial E) \subset (\hat{P}, \text{int } P)$ be a disk which contains the edges of all faces in $\Sigma$. It is then possible to choose the collection $\Sigma$ so that every disk in it lies locally on the same side of $E$ (cf. [6, §4]). Subject to these constraints, we further assume that $|\Sigma|$ is as small as possible (where $|\ldots|$ stands for cardinality or number of connected components).

Now, the 2-sphere $\hat{P}$ splits the filling solid torus of $X(r)$ into a finite collection of ‘1-handles’ $\mathcal{H}$. Following [6, §4], let $N(E, \Sigma) \subset X(r)$ be the regular neighborhood of the set $E \cup \mathcal{H}^* \cup \Sigma$, where $\mathcal{H}^* \subset \mathcal{H}$ denotes the
collection of 1-handles each of which is intersected by the boundary of at least one disk in $\Sigma$. Then $M$ can be taken to be the manifold $\hat{N}(E, \Sigma)$ (where $\hat{Y}$ is obtained by capping off each boundary 2-sphere component of $Y$ with a 3-ball); that $M$ has nontrivial torsion in first homology follows from [13]. We remark that we can take $E = \hat{P}$ whenever $\hat{P}$ separates, in which case $\partial N(E, \Sigma)$ consists of two 2-spheres, one of which is parallel to $\hat{P}$ and the other having fewer than $|\partial P|$ boundary components. We call the collection $\Sigma$ a generalized Scharlemann cycle, and say that $M$ has been obtained by attaching the generalized Scharlemann cycle $\Sigma$ to $E$ (or $\hat{P}$, as the case may be). Perhaps the simplest example of this situation is provided by the example of Gordon and Litherland [7, Appendix] (see also [15, §4]).

The construction of the manifold $M$ can be abstracted as follows: Throughout, we work in the PL-category, and all manifolds are assumed to be compact and orientable. Some familiarity of the reader with the papers [4, Chapter 2] and [8, 6] is assumed; we also refer the reader to [3] for standard definitions and notation about Heegaard splittings. Let $H_n$ be a genus $n$ handlebody with a fixed complete disk system $D = \{D_1, \ldots, D_n\}$; we call the circles $\partial D = \{\partial D_1, \ldots, \partial D_n\}$ the meridians of $H_n$. A collection of disjoint circles $C$ embedded in $\partial H_n$ is said to intersect $\partial D$ coherently if, for each $c \in C$ and $D \in D$, the intersection $c \cap \partial D$ is transverse and $|c \cdot D| = |c \cap D|$. This last condition can be restated as saying that each curve in $C$ intersects a given meridian circle of $H_n$ always in the same direction.

Now fix orientations of $H_n$ and the components of $\partial D$ and $C$. For each circle $c$ in $C$, let $a(c) = (a(c)_1, \ldots, a(c)_n)$ be the ordered $n$-tuple whose $i$-th entry is given by the algebraic intersection number $c \cdot D_i$ of $c$ with $\partial D_i$. The collection $C$ represents all types (relative to the meridians of $H_n$) if the set of integral vectors $\{a(c), c \in C\}$ represents all $n$-types in the sense of Parry [13], that is, if for any real vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ there is a vector $a(c) \in C$ such that all the nonzero terms in the expansion of the standard inner product $\langle a(c), x \rangle$ have the same sign.

For any collection $C$ of disjoint circles embedded in $\partial H_n$ and satisfying the conditions:

1) $C$ intersects $\partial D$ coherently,
2) $C$ represents all types,
3) no proper subcollection of $C$ represents all types, and
4) every meridian circle in $\partial D$ is intersected nontrivially by at least one circle in $C$,

the 3-manifold $M$ obtained by attaching 2-handles to $\partial H_n$ along the circles in $C$ and capping off any resulting 2-sphere boundary components will be called a $t$-manifold. The particular data $S = (H_n, D, C)$ used to construct $M$ will be called a genus $n$ $t$-manifold structure of $M$, and we will refer to
the curves in \( C \) as the *attaching circles* of the t-manifold structure. Observe that the integral matrix \( A_M \) whose rows are the vectors \( \{ a(c), c \in C \} \) is a presentation matrix for \( H_1(M) \).

It follows from the definition that any genus one t-manifold is either \( S^3 \) or a lens space. In the first part of this paper we give a complete classification of the 3-manifolds that admit a genus two t-manifold structure; this is the content of our first theorem:

**Theorem 1.1.** If \( M \) is a t-manifold with a genus two t-manifold structure, then \( M = M_1 \cup \partial M_2 \), where each \( M_i \) is a Seifert fibered space over a disk with at most two singular fibers, such that the regular fibers of \( M_1, M_2 \) intersect in one point in \( \partial M_1 = \partial M_2 \).

In particular, \( M \) is either toroidal, a Seifert fibered space over a 2-sphere with 3 singular fibers, a lens space, or \( S^3 \), and it is irreducible but not always Haken.

It is possible to determine when a 3-manifold \( M \) admitting a decomposition of the form \( M = M_1 \cup \partial M_2 \) as in Theorem 1.1 is indeed a t-manifold, via coherency invariants which are derived from the Euler numbers of \( M_1 \) and \( M_2 \). These invariants are constructed via a function \([\ldots] : \mathbb{R} \to \mathbb{Z}/2\mathbb{Z}\) defined as follows: \([x] = [x] + 1/2\) for each real number \( x \) that is not an integer (here \([\ldots]\) denotes the greatest integer function), and \([x] = x\) for each integer \( x \). Observe that \([\ldots]\) is an odd periodic function of period 1.

**Theorem 1.2.**

(a) Let \( M \) be a 3-manifold of the form \( M = M \cup \partial M' \), where \( M, M' \) are Seifert-fibered spaces over a disk with two singular fibers, glued along their torus boundaries in such a way that their regular fibers intersect in one point. Let the Seifert invariants of \( M \) and \( M' \) be \( (a_1, p_1; a_2, p_2) \) and \( (a'_1, p'_1; a'_2, p'_2) \), where \( a_i, a'_i > 1 \) for each \( i \). Then \( M \) has a genus two t-manifold structure iff \( ([p_1/a_1] + [p_2/a_2]) \cdot ([p'_1/a'_1] + [p'_2/a'_2]) < 0 \).

(b) Let \( M \) be a Seifert fibered manifold over the 2-sphere with three singular fibers and Seifert invariant \( (a_1, p_1; a_2, p_2; a_3, p_3) \), where \( a_i > 1 \) for each \( i \). Then \( M \) admits a genus two t-manifold structure iff \([p_1/a_1] + [p_2/a_2] + [p_3/a_3] \neq \pm 1/2\).

Before discussing higher genus t-manifold structures, we slightly generalize the notion of types representation. We will say that a nonempty set of vectors \( X \) in \( \mathbb{R}^n \) represents all types if for each vector \( v \in \mathbb{R}^n \) there is a vector \( x \in X \) such that all the nonzero terms in the expansion of the standard inner product \( \langle v, x \rangle \) have the same sign. Observe that we allow sets of nonintegral vectors to represent all types. We extend this definition to matrices, and say that a real \( k \times n \) matrix represents all types if its set of row vectors, viewed as elements of \( \mathbb{R}^n \), represent all types. The following result is the starting point for our analysis of the topological structure of
t-manifolds; it essentially follows from Lemma 4.4 in [9] (see also [6, Lemma 3.2]), though our approach to the proof is somewhat different:

**Theorem 1.3.** Let $A$ be a $k \times n$ matrix representing all types. Then one of the following must hold:

1. $A$ has a column of zeroes;
2. A proper subset of the rows of $A$ represents all types;
3. $A$ has rank $n$.

Matrices which represent all types and do not satisfy conditions (a) or (b) of Theorem 1.3 will be called *proper matrices*. In this sense, a finite set of vectors in $\mathbb{R}^n$ which represents all types ‘efficiently’ must be a spanning set, and proper matrices are those matrices which represent all types ‘efficiently’. Observe that, for a t-manifold $M$ with t-manifold structure $S = (H_n, D, C)$, the associated integral matrix $A_M$ is proper. The following result summarizes some immediate properties of $M$ and the matrix $A_M$ (cf. [6, Theorem 4.3]):

**Theorem 1.4.** Let $M$ be any t-manifold with a genus $n$ t-manifold structure and t-presentation matrix $A_M \neq (1)$. Then:

1. $A_M$ has rank $n$;
2. $M$ is closed;
3. $H_1(M)$ is finite and nontrivial.

It is possible to extract some more information about the topology of a t-manifold by studying the distribution of zero entries in its t-presentation matrices. Define the *girth* of any $k \times l$ matrix as the number $k + l$. The *zero girth* (0-girth for short) of a matrix is then defined as the largest girth of any of its zero submatrices. The following result gives a bound on the 0-girth of a square matrix that represents all types properly:

**Theorem 1.5.** An $n \times n$ matrix $A$ that represents all types is proper iff its 0-girth is at most $n - 1$.

This result can be used to give a nice characterization of proper square matrices (see Corollary 5.7) and, along with well-known results of A. Casson and C. Gordon on reducible Heegaard splittings of 3-manifolds, to show that some t-manifolds having ‘square’ t-manifold structures of smallest possible genus may not be hyperbolic.

**Theorem 1.6.** Let $M$ be a t-manifold with a genus $n$ t-manifold structure and square t-presentation matrix $A_M$. If the Heegaard genus of $M$ is $n$, then either:

1. $M$ is reducible or contains an incompressible embedded torus, or
2. the 0-girth of $A_M$ is at most $n - 2$. 

Moreover, if the matrix $A_M$ has a $2 \times 2$ submatrix which does not represent all types, then (a) holds.

A t-manifold can be thought of as a generalized lens space, just like a generalized Scharlemann cycle generalizes the notion of Scharlemann cycle. From this point of view, it may not be surprising that no t-manifold with a Heegaard genus two t-manifold structure is hyperbolic. One can then ask whether there are hyperbolic t-manifolds at all. The above theorem gives a partial answer to this question in the negative, but the general case is still open.

We also prove that t-manifolds arise in a natural way from Dehn fillings of knot spaces along different planar boundary slopes; the following result is a mild extension of [15, Theorem 1.3]:

**Theorem 1.7.** Let $X$ denote a knot space with distinct planar boundary slopes $r, s$. Then, either $X(s)$ has a lens space (t-manifold, resp.) connected summand or $X(r)$ has a t-manifold (lens space, resp.) connected summand and at most one prime factor of $X(r)$ is not part of some t-manifold (it is not a lens space, resp.) connected summand.

The paper is organized as follows: We prove in Section 2 some basic results on systems of arcs in 4-punctured 2-spheres, which will be used in Section 3 to prove Theorem 1.1. Theorem 1.2 is proved in Section 4, where coherency invariants of the Seifert fiber structures are introduced. Theorem 1.3 and Theorem 1.5 are proved in Section 5, along with other properties of proper matrices and their 0-girths. Proofs of Theorems 1.4, 1.6, and 1.7 are given in Section 6.

I want to thank Andrew Casson for many suggestions and helpful discussions on these results, some of which formed part of my thesis.

**2. Systems of arcs in 4-punctured 2-spheres.**

Let $H$ be an orientable genus two handlebody with fixed meridian disks $D, \overline{D}$ and corresponding meridian circles $m = \partial D, \overline{m} = \partial \overline{D}$, and let $\gamma$ be an essential closed 1-submanifold of $\partial H$ that intersects $m \cup \overline{m}$ transversely. Cutting $\partial H$ along the meridian circles $m, \overline{m}$ yields a 4-punctured sphere $S_0$; denote the punctures corresponding to $m$ by $m_1, m_2$, and those corresponding to $\overline{m}$ by $\overline{m}_1, \overline{m}_2$. We say that the collection $\Gamma(\gamma) = \gamma \cap S_0 \subset S_0$ is **standard** if its components are essential arcs in $S_0$ which can be sorted out into six subcollections $\Gamma_i(\gamma)$, $1 \leq i \leq 6$, as follows:

- $\Gamma_1(\gamma)$: Arcs connecting $m_1$ and $\overline{m}_1$,
- $\Gamma_2(\gamma)$: Arcs connecting $m_2$ and $\overline{m}_2$,
- $\Gamma_3(\gamma)$: Arcs connecting $m_1$ and $\overline{m}_2$,
- $\Gamma_4(\gamma)$: Arcs connecting $m_2$ and $\overline{m}_1$,
- $\Gamma_5(\gamma)$: Arcs connecting $m_1$ and $m_2$,
- $\Gamma_6(\gamma)$: Arcs connecting $\overline{m}_1$ and $\overline{m}_2$.

Observe that $\Gamma(\gamma)$ does not contain any arc connecting a boundary component of $S_0$ to itself (see Figure 1) and has no circle components. We will
omit the argument $\gamma$ when this does not cause any confusion. Two arcs $a, b$ of $\Gamma(\gamma)$ are parallel in $S_0$ if the closure of one of the components of $S_0 \setminus (a \cup b)$ is a rectangle.

**Lemma 2.1.** Let $\Gamma(\gamma)$ be a standard system of arcs in $S_0$.

(a) If $\Gamma_5 \neq \emptyset$ ($\Gamma_6 \neq \emptyset$, resp.), then either:
   (i) Each collection $\Gamma_5$ and $\Gamma_6$ consists of mutually parallel arcs, or
   (ii) $\Gamma_5$ ($\Gamma_6$, resp.) contains two non-parallel arcs $a, b$ such that every arc in $\Gamma_5$ ($\Gamma_6$, resp.) is parallel to one of $a$ or $b$, and $\Gamma_6$ ($\Gamma_5$, resp.) is empty.

(b) Let $n_i = |\Gamma_i|$ for each $i$. Then $n_1 = n_2$ and $n_3 = n_4$.

**Proof.** Suppose that $\Gamma_5$ contains two non-parallel arcs $a, b$. Cut $S_0$ along $a$ to obtain a 3-punctured sphere $S_0'$ with boundary components $\overline{m}_1, \overline{m}_2, C$ and
containing \( b \), where \( C \) consists of one arc from each of \( m_1 \) and \( m_2 \) connected by two arcs parallel to \( a \), as indicated in Figure 2(a).

If there is an arc \( c \) in \( S'_0 \) disjoint from \( b \) connecting \( \overline{m}_1 \) and \( \overline{m}_2 \), then cutting \( S'_0 \) along \( c \) yields an annulus \( A \) with two boundary components \( C, C' \) and containing \( b \). Since both endpoints of \( b \) lie on \( C \), \( b \) must be boundary-parallel to \( C \) in \( A \) and hence it must be parallel to \( a \) in \( S_0 \), which is a contradiction. Therefore, \( \Gamma_6 \) is empty and \( b \) separates \( \overline{m}_1 \) from \( \overline{m}_2 \) in \( S'_0 \) (Figure 2(b)).

Cut \( S'_0 \) along \( b \) to obtain two annuli \( A_1, A_2 \), each containing \( m_1, m_2 \), respectively. Then any arc in \( A_1 \) or \( A_2 \) connecting \( m_1 \) and \( m_2 \) must be boundary-parallel (see Figure 2(b)), hence parallel in \( S_0 \) to one of \( a \) or \( b \).

Suppose now that all the arcs of \( \Gamma_5 \) are parallel. By amalgamating all these arcs if necessary, we may assume that \( \Gamma_5 \) contains only one arc \( d \). Since \( d \) does not separate \( m_1 \) from \( m_2 \), \( \Gamma_6 \) may not be empty. If indeed \( \Gamma_6 \) is not empty, then all its arcs must be mutually parallel, for otherwise the argument above would apply to show that \( \Gamma_5 \), contrary to hypothesis, is empty. This establishes Part (a).

For Part (b), observe that the number of endpoints of arcs of \( \Gamma \) in \( m_1 \) and \( m_2 \) match, and so
\[
 n_1 + n_3 + n_5 = n_2 + n_4 + n_5,
\]
which implies that \( n_1 + n_3 = n_2 + n_4 \). Similarly, comparing the number of endpoints of \( \Gamma \) in \( \overline{m}_1 \) and \( \overline{m}_2 \) yields \( n_1 + n_4 = n_2 + n_3 \), whence \( n_1 = n_2 = n_3 = n_4 \).

We call a standard system of arcs \( \Gamma(\gamma) \subset S_0 \) split whenever the arcs in \( \Gamma_5(\gamma) \) or \( \Gamma_6(\gamma) \) split into two nonempty, non-parallel collections of parallel arcs as in Lemma 2.1 (a)(ii). Otherwise, \( \Gamma(\gamma) \) is said to be non-split.

A waist circle of \( \partial H \) is an essential circle \( w \subset \partial H \) separating \( m \) from \( \overline{m} \); such a circle always bounds a disk properly embedded in \( H \). Any waist circle is said to be \( \Gamma(\gamma) \)-simple if it is transverse to \( \Gamma(\gamma) \) and disjoint from \( \Gamma_5(\gamma) \cup \Gamma_6(\gamma) \).

**Lemma 2.2.** There exists a \( \Gamma \)-simple waist circle \( w \) iff \( \Gamma \) is non-split. Such a waist circle is unique up to isotopy if \( \Gamma_5 \cup \Gamma_6 \) is not empty, and may be assumed to intersect each arc of \( \Gamma \setminus (\Gamma_5 \cup \Gamma_6) \) in one point.

**Proof.** Let \( w \) be a \( \Gamma \)-simple waist circle. Without loss of generality, we may assume \( \Gamma_5 \) is not empty. If \( \Gamma_5 \) contains two arcs \( a, a' \), then the component of \( S_0 \) cut along \( w \cup a \) which contains \( a' \) is an annulus, with one boundary component consisting of one arc from each of the circles \( m_1, m_2 \) connected by two arcs parallel to \( a \), while the other boundary component is a curve parallel to \( w \). It is then clear that \( a' \) must be parallel to \( a \), so \( \Gamma \) is non-split.

Conversely, suppose \( \Gamma \) is non-split. Without loss of generality, we may assume that \( \Gamma_5 \) contains exactly one arc (either by creating it or by amalgamating all its arcs). Denote such an arc by \( b \), and let \( N \) be a regular
neighborhood of \( b \) in \( S_0 \), small enough to be disjoint from all other arcs in \( \Gamma \). We then obtain a \( \Gamma \)-simple waist circle \( w \) by band-connecting \( m_1 \) and \( m_2 \) along \( N \). Observe that each arc of \( \Gamma \setminus (\Gamma_5 \cup \Gamma_6) \) is intersected by \( w \) exactly in one point.

Now suppose \( w \) is a \( \Gamma \)-simple waist in \( S_0 \) and that \( \Gamma_5 \) contains an arc \( c \). Cut \( S_0 \) along the arc \( c \) to obtain a 3-punctured sphere \( S'_0 \) with boundary components \( \overline{m}_1, \overline{m}_2, C \), where \( C \) consists of pieces from \( m_1, m_2 \) and two arcs parallel to \( c \). Since \( w \) separates \( \overline{m}_1 \cup \overline{m}_2 \) from \( C \), \( w \) and \( C \) cobound an annulus in \( S'_0 \) and are therefore parallel in \( S_0 \). The uniqueness of \( w \) up to isotopy follows. \( \square \)

### 3. Genus two t-manifold structures.

This section is devoted to the proof of Theorem 1.1. The properties of standard systems of arcs established in the previous section will enable us to get a detailed picture of genus two t-manifold structures.

Clearly, a genus two t-manifold structure on a 3-manifold \( \mathcal{M} \) has exactly two attaching curves, and since its first homology is finite, \( \mathcal{M} \) is a closed manifold. Denote the meridian circles of the genus two handlebody \( H \) by \( u, v \), and assume that \( \mathcal{M} \) has a t-manifold structure consisting of two attaching circles \( x, y \) embedded in \( \partial H \). Hence \( H' = \overline{M \setminus H} \) is a handlebody and so the pair \((H, H')\) is a Heegaard splitting of \( \mathcal{M} \). Observe that the circles \( u, v \) in \( \partial H' \) also give a t-manifold structure to \( \mathcal{M} \) with respect to the meridians \( x, y \) of \( H' \).

We denote \( \partial H \) cut along \( u \cup v \) by \( S_0 \), and label its boundary components by \( u_1, u_2 \) and \( v_1, v_2 \). Similarly, we denote \( \partial H' \) cut along \( x \cup y \) by \( S'_0 \) and label its boundary components by \( x_1, x_2 \) and \( y_1, y_2 \).

**Lemma 3.1.**

(a) The collections of arcs \( \Gamma(x \cup y) \subset S_0 \) and \( \Gamma(u \cup v) \subset S'_0 \) are standard.

(b) Without loss of generality, we may assume that

\[
\begin{align*}
\Gamma_1(x \cup y) &\cup \Gamma_2(x \cup y) \subset x \quad \text{and} \quad \Gamma_3(x \cup y) \cup \Gamma_4(x \cup y) \subset y, \\
\Gamma_1(u \cup v) &\cup \Gamma_2(u \cup v) \subset u \quad \text{and} \quad \Gamma_3(u \cup v) \cup \Gamma_4(u \cup v) \subset v.
\end{align*}
\]

(c) Each of the collections \( \Gamma_i(x \cup y) \), \( \Gamma_i(u \cup v) \) is nonempty for \( 1 \leq i \leq 4 \) and consists of parallel arcs.

**Proof.** From the definition of t-manifold, no arc in \( \Gamma(x \cup y) \) connects a boundary component of \( S_0 \) to itself; hence \( \Gamma(x \cup y) \) is standard, so (a) follows.

Suppose now that an arc of \( \Gamma_1(x \cup y) \) is part of \( x \), oriented to run from \( u_1 \) to \( v_1 \). Then \( x \) must intersect \( u \) always from \( u_2 \) to \( u_1 \), and \( v \) from \( v_1 \) to \( v_2 \). By coherency and the fact that \( x \) and \( y \) must represent all types, it follows that \( \Gamma_1(x \cup y) \cup \Gamma_2(x \cup y) \subset x \) and \( \Gamma_3(x \cup y) \cup \Gamma_4(x \cup y) \subset y \). The other cases are similar, so (b) holds.
If, say, $\Gamma_1(x \cup y)$ is empty, then $x$ must be disjoint from one of $u$ or $v$, violating condition 4) in the definition of $t$-manifold. The fact that the collections in (c) consist of parallel arcs follows now from Lemma 2.1.

Let $n_i = |\Gamma_i(x \cup y)|$ and $n'_i = |\Gamma_i(u \cup v)|$. From Lemma 2.1(b), $n_i = n_j$ and $n'_i = n'_j$ for $\{i, j\} = \{1, 2\}, \{3, 4\}$. For convenience, we will use the notation:

\[ a = n_1, \quad b = n_3, \quad \alpha = n'_1, \quad \beta = n'_3. \]

The proof of Theorem 1.1 now splits into two cases:

**Case 1.** Both $\Gamma(x \cup y) \subset S_0$ and $\Gamma(u \cup v) \subset S'_0$ are split.

**Case 2.** Either $\Gamma(x \cup y) \subset S_0$ or $\Gamma(u \cup v) \subset S'_0$ is non-split.

**Proof of Case 1.** Suppose that $\Gamma_5(x \cup y)$ splits, so that $\Gamma_6(x \cup y)$ is empty by Lemma 2.1(a)(ii). Hence, the endpoints of $\Gamma(x \cup y)$ in $v_1$ are just those of $\Gamma_1(x \cup y)$ and $\Gamma_4(x \cup y)$. Since $\Gamma_1(x \cup y) \subset x$ and $\Gamma_4(x \cup y) \subset y$, $x \cup y$ must intersect $v$ in the pattern of Figure 3. Therefore, each of $\Gamma_3(u \cup v)$ and $\Gamma_4(u \cup v)$ must consist of exactly one arc, so $\beta = 1$. If now, say, $\Gamma_5(u \cup v) \subset S'_0$ splits, so $\Gamma_6(u \cup v)$ is empty by Lemma 2.1(a)(ii), then

\[ |v \cap y| = |v_1 \cap y| = |\Gamma_6(u \cup v) \cap v| + |\Gamma_4(u \cup v)| = 1 \]

and so $v$ and $y$ intersect in one point. It follows that the Heegaard splitting $(H, H')$ of $M$ can be reduced to a genus one Heegaard splitting; hence $M$ is a lens space and has the form required in the theorem.

**Proof of Case 2.** Suppose that $\Gamma(x \cup y)$ is non-split. Then Lemma 2.2 guarantees the existence of a $\Gamma(x \cup y)$-simple waist circle $w$ of $H$. We proceed according to two subcases:

**Subcase 2(a).** Either $a = 1$ or $b = 1$.

Assume, without loss of generality, that $a = 1$. Let $D_w \subset H$, $D_x \subset H'$ be properly embedded disks bounded by $w$ and $x$, respectively. By construction (see Lemma 2.2), $w$ intersects each arc of $\Gamma_1(x \cup y) \cup \Gamma_2(x \cup y) \subset x$ in one point and hence (since $a = n_1 = n_2 = 1$) $x$ in two points.
Let \( \zeta \) be a properly embedded arc in \( D_x \) connecting these two points, and let \( N_\zeta \) be a small regular neighborhood of \( \zeta \) in \( H' \). The genus three handlebody \( H_\zeta = H \cup N_\zeta \) contains a properly embedded annulus \( A_\zeta = D_w \cup B \), where \( B \) is a properly embedded band in \( N_\zeta \), that separates \( H_\zeta \) into a pair of genus two handlebodies \( H_{\zeta,u}, H_{\zeta,v} \) containing \( u, v \), respectively. Observe that \( \zeta \) can be extended to a core of the annulus \( A_\zeta \) via an arc in \( D_w \) (see Figure 4).

Now, \( N_\zeta \) separates \( D_x \) into two disks \( D_{x,u} \) and \( D_{x,v} \), properly embedded in \( H_{\zeta,u} \) and \( H_{\zeta,v} \), respectively; since these two disks intersect the meridians of the \( N_\zeta \)-handles of \( H_{\zeta,u}, H_{\zeta,v} \) in one point, respectively, it follows that

\[
H \cup N(D_x) = (H_{\zeta,u} \cup N(D_{x,u})) \cup_{A_\zeta} (H_{\zeta,v} \cup N(D_{x,v})) = V_{\zeta,u} \cup_{A_\zeta} V_{\zeta,v},
\]

where \( V_{\zeta,u}, V_{\zeta,v} \) are solid tori whose meridian circles intersect the core of \( A_\zeta \) in \( s = |x \cap u| \) and \( t = |x \cap v| \) points, respectively. Hence, \( N(H \cup D_x) \) is a Seifert fibered space over a disk with (at most) two singular fibers of indices \( s \) and \( t \).

It only remains to attach a 2-handle to \( H \cup N(D_x) \) along \( y \) and cap off the resulting sphere component to obtain \( M \). Observe that \( y \) intersects each boundary component of \( A_\zeta \) in \( b \) points (recall \( b = n_3 = n_4 \), see Figure 4). Therefore, \( M \) is a Seifert fibered space over the 2-sphere with (at most) three singular fibers of indices \( s, t \) and \( b \). Replacing the fibration in a fibered solid torus neighborhood of some singular fiber in a suitable way shows that \( M \) has the required form.

**Subcase 2(b).** \( a > 1 \) and \( b > 1 \).
If $\Gamma(u \cup v) \subset S'_0$ splits then $a = 1$ or $b = 1$ according to the first half of the argument for Case 1 above. Therefore $\Gamma(u \cup v)$ is non-split; by Subcase 2(a), we may also assume that $\alpha > 1$ and $\beta > 1$.

By construction (see Lemma 2.2), $w$ intersects $x \cup y$ in the pattern given in Figure 5, from which we deduce that:

1. $\Gamma(w) \subset S'_0$ is standard: For no arc of $\Gamma(w)$ in $S'_0$ runs from a component of $\partial S'_0$ to itself.
2. $|\Gamma_i(w)| = 1$ for $1 \leq i \leq 4$.
3. $|w \cap x| = 2 + |\Gamma_5(w)|$ and $|w \cap y| = 2 + |\Gamma_6(w)|$: To see this, split $w$ at the four points A, B, C, D marked in Figure 5. The arc $AB$ contains a subset $\Gamma'_5(w)$ of the arcs in $\Gamma_5(w)$, while the arc $CD$ contains the remaining arcs $\Gamma''_5(w)$. Hence, $|AB \cap x| = 1 + |\Gamma'_5(w)|$ and $|CD \cap x| = 1 + |\Gamma''_5(w)|$, so that $|w \cap x| = 2 + |\Gamma_5(w)|$ holds. The other equality follows in a similar way.
4. $|w \cap x| = 2a$ and $|w \cap y| = 2b$: This follows from the construction of $w$ in Lemma 2.2.
5. $\Gamma(w) \subset S'_0$ is non-split: For suppose that $\Gamma_5(w) \subset S'_0$ splits, so that $\Gamma_6(w)$ is empty. Comparing items (3) and (4) above yields $2 = 2b$, hence $b = 1$, contrary to hypothesis.

These facts imply that $\Gamma(u \cup v \cup w) \subset S'_0$, which is automatically standard, is also non-split; a similar remark holds for the collection $\Gamma(x \cup y \cup w') \subset S'_0$. Let $w'$ be a $\Gamma(u \cup v \cup w)$-simple waist circle which bounds a properly embedded disk $D_{w'}$ in $H'$.

**Claim 1.** $|w \cap w'| = 4$ and $w$ is a $\Gamma(x \cup y \cup w')$-simple waist.

*Proof.* Since $w'$ is a $\Gamma(w)$-simple waist of $H'$, then $w'$ intersects $w$ only at the arcs $\Gamma_i(w)$ for $i = 1, \ldots, 4$. By item (2) above, we know that $|\Gamma_i(w)| = 1$ for $i = 1, \ldots, 4$. Hence,

$$|w \cap w'| = |\Gamma_1(w)| + |\Gamma_2(w)| + |\Gamma_3(w)| + |\Gamma_4(w)| = 4.$$ 

In particular, this implies that $w$ is a $\Gamma(x \cup y \cup w')$-simple waist. \qed
Claim 2. The circles \( w \cup w' \) split the surface \( S = \partial H = \partial H' \) into two bands (rectangles) \( B, B' \) and two annuli \( A, A' \).

Proof. Each of the circles \( w, w' \) separates \( S \) into two once-punctured tori. Let \( T_w \) be the closure of one of the components of \( S \setminus w \), so that \( T_w \) is a punctured torus with boundary \( w \). Since \( w' \) separates \( S \), the two arcs of \( w' \cap T_w \) must be parallel in \( T_w \) and therefore separate \( T_w \) into a band \( B \) and an annulus \( A \). Similarly, \( w' \) separates the other component of \( S \setminus w \) into a band \( B' \) and an annulus \( A' \) (see Figure 6).

The disk \( D_w \) separates \( H \) into two components whose closures are solid tori \( V_u, V_v \), with meridian circles \( u \subset \partial V_u \) and \( v \subset \partial V_u \). Without loss of generality, we may assume that \( \partial V_u = A \cup B \cup D_w \) and \( \partial V_v = A' \cup B' \cup D_w \). Similarly, the waist disk \( D_w' \) separates \( H' \) into two components whose closures are solid tori \( V_x, V_y \), with meridian circles \( x \subset \partial V_x \) and \( y \subset \partial V_y \). Here, we may assume that \( \partial V_x = A \cup B \cup D_w \) and \( \partial V_y = A' \cup B' \cup D_w' \).

Now let \( T \) be the torus \( D_w \cup B \cup B' \cup D_w' \), which is embedded in \( M \). It follows from the previous paragraph (see also Figure 6) that \( T \) separates \( M \) into two components whose closures are \( V_u \cup A \cup V_x \) and \( V_v \cup A' \cup V_y \). Since each arc of \( w \) in \( \partial V_x \) intersects \( x \) geometrically in \( a \) points (see Figure 5), the core of \( A \) must intersect \( x \) geometrically in \( a \) points. Similarly, the core of \( A' \) intersects \( y \) and \( v \) geometrically in \( b \) and \( \beta \) points, respectively.

The fibration of \( A \) by core circles can be easily extended to a Seifert fibration of \( V_u \) and \( V_x \), with singular fibers of indices \( a \) and \( \alpha \), respectively, while the circle fibration of \( A' \) extends to a Seifert fibration of \( V_v \) and \( V_y \) with singular fibers of indices \( b \) and \( \beta \), respectively. Hence, the manifolds
$M_1 = V_u \cup_A V_x$ and $M_2 = V_v \cup_A V_y$ are Seifert fibered spaces over a disk with two singular fibers of indices $a, \alpha$ and $b, \beta$, respectively.

To determine how the fibers of $M_1$ and $M_2$ intersect at the boundary, we observe that:

- $V_x$ and $V_y$ lie on opposite sides of $T$ and intersect at $D_{w'}$, and
- the annulus complementary to $A$ in $\partial V_x$ is given by $B \cup D_{w'}$, while the annulus complementary to $A'$ in $\partial V_y$ is given by $B' \cup D_{w'}$.

Now, it is clear that the cores of the annuli $B \cup D_{w'} \subset \partial M_1$ and $B' \cup D_{w'} \subset \partial M_2$ intersect in one point; in fact, the disk $D_{w'}$ can be thought of as being a fat regular neighborhood of such a point of intersection; it follows that the fibers of $M_1$ and $M_2$ intersect in one point. The proof of Theorem 1.1 is now complete. □


In this section we prove Part (a) of Theorem 1.2; Part (b) is somewhat similar, and we refer the reader to [14] for a detailed proof. The determination of when a decomposition of a manifold $M$ as in Theorem 1.1 gives rise to a $t$-manifold structure will depend on invariants of the Seifert fibered structures that detect the necessary coherent intersections. We begin by describing how these invariants are constructed via Seifert invariants. The set of notes by Jankins and Neumann [12] contain a nice exposition of basic facts about Seifert fibered spaces, some of which we present here to set the notation we will be using in the sequel.

Let $p : M \to D$ be an oriented Seifert fibered space over a disk with two singular fibers of indices $a_1, a_2$ and fixed orientation on all fibers. Denote by $x_i$ the projection point in $D$ of the singular fiber of index $a_i$.

Let $\alpha \subset D$ be a properly embedded arc separating $D$ into two disks $D_1, D_2$ with $x_i \in \text{int } D_i$. The fibered annulus $A = p^{-1}(\alpha)$ then separates $M$ into two solid tori $V_1$ and $V_2$, with $p^{-1}(x_i) \subset V_i$. We take one of the components of $\partial A$ as representative of the regular fibers of $M$, and denote it by $h$. Let $D^* = D \setminus \text{int}(D'_1 \cup D'_2)$, where $D'_i \subset D_i$ is a small disk containing $x_i$ in its interior, and let $M^* = p^{-1}(D^*)$ be the associated trivial $S^1$-bundle over $D^*$.

Assume now that an oriented simple closed curve $\mu \subset \partial M$ intersecting the fibers of $\partial M$ transversely in one point is given. Any section $s : D^* \to M^*$ of $p|M^*$ such that $s(\partial D) = \mu$ can be used to frame the torus $\partial V_i$ via the curves $\mu_i = s(\partial D_i)$ (framing meridians) and $h$ (longitude), where each $\mu_i$ inherits its orientation from $\mu$; we call $\mu_i$ a $\mu$-meridian of $V_i$. Observe that $\mu_1 \cap \mu_2$ is an arc in $A$ receiving opposite orientations from the framing meridians $\mu_1, \mu_2$. To stress the dependency between the circle $\mu$ and the section $s$ used to frame $M$, we refer to such a framing as a $\mu$-framing of $M$. 


With respect to a particular \( \mu \)-framing, the meridian circle \( m_i \subset \partial V_i \) of \( V_i \) can be oriented so that, in \( H_1(\partial V_i) \), it is of the form
\[
m_i = a_i \mu_i + p_i h, \quad a_i > 1.
\]
The ordered 4-tuple \((a_1, p_1; a_2, p_2)\) is called the Seifert invariant of \( M \). The Euler number of the Seifert fibration on \( M \) is defined as \( e(M) = a_1/p_1 + a_2/p_2 \). These invariants are sensitive to changes in the orientation of the manifold and/or the fibers, e.g., \( e(-M) = -e(M) \), where \(-M\) denotes the manifold \( M \) with the opposite orientation.

**Remark 4.1.** The Seifert invariant also depends on the particular section used to construct it. In fact, if a different section is used, then the new Seifert invariant obtained will be of the form \((a_1, p_1 + k_1 a_1; a_2, p_2 + k_2 a_2)\), where \( k_1 \) and \( k_2 \) are integers such that \( k_1 + k_2 = 0 \) (see [12]). The Euler number of \( M \), however, is independent of the section used to get the Seifert invariant.

The function \([\ldots]\) defined in the Introduction can now be used to construct an invariant of \( M \) similar to its Euler number. If the Seifert invariant of \( M \) with respect to some \( \mu \)-framing is \((a_1, p_1; a_2, p_2)\), define
\[
c(M) = [p_1/a_1] + [p_2/a_2].
\]
The periodicity of the function \([\ldots]\) and Remark 4.1 immediately imply that \( c(M) \) is an invariant of \( M \); that is, like \( e(M) \), \( c(M) \) depends only on the oriented Seifert fiber structure of \( M \) and not on the particular Seifert invariant used in its construction.

Let \( M, M' \) be oriented Seifert fibered spaces over a disk with two singular fibers, and suppose that \( \mathcal{M} = M \cup_M M' \) in such a way that the fibers of \( M, M' \) at the boundary intersect in one point *after an orientation reversing gluing*. Under these conditions, the manifold \( \mathcal{M} \) admits certain Heegaard splittings similar to the ones considered in the proof of Subcase 2(b) of Theorem 1.1 (Section 3).

Let \( h \subset \partial M \), \( h' \subset \partial M' \) be fibers intersecting in one point; we use these (-oriented) fibers to induce framings on \( M, M' \) in the sense of the previous section. That is, \( M \) is given an \( h' \)-framing and \( M' \) and \( h \)-framing, and we choose our notation so that:
\[
M = V_1 \cup_A V_2, \quad M' = V'_1 \cup_{A'} V'_2, \\
fiber = h \subset \partial A, \quad fiber = h' \subset \partial A', \\
\mu_i = h'-\text{meridian of } \partial V_i, \quad and \quad \mu_i' = h-\text{meridian of } \partial V_i', \\
m_i = a_i \mu_i + p_i h, \quad a_i > 1, \quad m_i' = a_i' \mu_i' + p_i' h', \quad a_i' > 1, \\
\mu_i \cdot h = 1 \text{ in } \partial V_i, \quad \mu_i' \cdot h' = 1 \text{ in } \partial V_i'.
\]
Seifert invariant = \((a_1, p_1; a_2, p_2)\) \quad Seifert invariant = \((a_1', p_1'; a_2', p_2')\).

Let \( T \) be the embedded separating torus \( \partial M = \partial M' \) in \( \mathcal{M} \). The four circles \( \partial A \cup \partial A' \) split \( T \) into four squares, each of which represents one of
the intersections $V_i \cap V_j'$ for $\{i, j\} \subset \{1, 2\}$, as in Figure 8. Clearly, $\mathcal{M}$ admits the following genus two Heegaard splittings:

(a) $\mathcal{M} = (V_1 \cup V_1 \cap V_1') \cup \partial (V_2 \cup V_2 \cap V_2')$

(b) $\mathcal{M} = (V_1 \cup V_1 \cap V_2') \cup \partial (V_2 \cup V_2 \cap V_1')$

We will call any genus two Heegaard splitting of $\mathcal{M}$ obtained in this way standard. Define the coherency invariant $C(\mathcal{M})$ of $\mathcal{M}$ as

$$C(\mathcal{M}) = c(\mathcal{M})c(\mathcal{M}') = ([p_1/a_1] + [p_2/a_2]) \cdot ([p_1'/a_1'] + [p_2'/a_2']).$$

Note that this invariant is independent of the orientations of $\mathcal{M}, \mathcal{M}'$.

**Proof of Theorem 1.2.** We will show that any standard Heegaard splitting of the manifold $\mathcal{M}$ gives rise to a t-manifold structure on $\mathcal{M}$ iff $C(\mathcal{M}) < 0$. To fix notation, we assume that the Heegaard splitting under consideration is of Type (a).

Let $H_1 = V_1 \cup V_1'$ and $H_2 = V_2 \cup V_2'$. We turn the meridian circles $m_1, m_1'$ of $V_1, V_1'$ into meridian circles of $H_1$ by isotoping them away from the square $V_1 \cap V_1'$; similarly, we isotope $m_2, m_2'$ in $H_2$ away from the square $V_2 \cap V_2'$. Figure 7 shows $m_1 \subset \partial V_1$ with all regions involved.

**Remark 4.2.** Observe that, as $m_1 = a_1\mu_1 + p_1 h'$ and $a_1 > 0$, the meridian circle $m_1$ and $\mu_1$ (which is essentially $h'$) both intersect $h$ in the same direction. A similar remark applies to the other meridians.

Consider now the genus two handlebody $H_1$ with meridian circles $m_1, m_1'$ and attaching curves $m_2, m_2' \subset \partial H_1$. In order for $\mathcal{M}$ to get a t-manifold structure from this Heegaard diagram, the intersections between the attaching curves and the meridian circles of $H_1$ must be coherent and represent all types. We first observe that:
Figure 8. $m'_1, m_2$ and $m_1, m'_2$ always intersect coherently.

Figure 9. The arcs of $m_1, m_2$ in the annulus $A$.

- All the intersections of $m_1$ with $m_2$ occur in the annulus $A$, while all the intersections of $m_1$ with $m'_2$ occur in the square $V_1 \cap V'_2$, and
- all the intersections of $m'_1$ with $m'_2$ occur in the annulus $A'$, while all the intersections of $m'_1$ with $m_2$ occur in the square $V'_1 \cap V_2$.

From the way the meridians intersect the square regions, it follows that such intersections must necessarily be coherent, as in Figure 8. The orientations of the arcs shown in Figure 8 were obtained from our observations in Remark 4.2. Hence, the signs of the intersections within the squares depend only on the framings induced by $h \cup h'$ on $M, M'$.

The intersections occurring in the annuli $A$ and $A'$ are a bit more involved. Figure 9 shows how $m_1$ and $m_2$ may intersect in $A \subset V_1$. Observe that,
since \( m_1 \) and \( m_2 \) do not intersect the squares \( V_1 \cap V_1' \) and \( V_2 \cap V_2' \), they can only intersect \( A \) in a restricted way which may give rise to intersections of opposite signs (i.e., incoherent intersections).

To study this coherency problem, we first turn \( A \) into a torus \( T_A \) by gluing the boundary circles of \( A \) in such a way that the arcs \( \mu_1 \cap A, m_1 \cap A, \) and \( m_2 \cap A \) each give rise to a closed curve in \( T_A \). (Equivalently, we collapse the annulus in \( \partial V_1 \) complementary to \( A \) onto \( h \) to get \( T_A \).) We call the resulting curves \( \mu_1|A, m_1|A, m_2|A, \) respectively, and orient them following the orientation of the arcs used in their construction. As framing for \( T_A \), we take the circles \( \mu_1|\circ A \subset T_A \) and \( h \subset T_A \) with \( \mu_1|A \cdot h = 1 \). In this way, since the framed tori \( T_A \) and \( \partial V_1 \) are essentially ‘the same’, we have that

\[
m_1|A = a_1\mu_1|A + p_1h,
\]

and since the two arcs \( \mu_1 \cap A \) and \( \mu_2 \cap A \) agree but have opposite orientations in \( A \) (see Figure 9), then

\[
m_2|A = -a_2\mu_1|A + p_2h.
\]

We further observe that \( h \subset T_A \) naturally splits into two arcs \( h_1, h_2 \), such that \( m_1|A \) intersects \( h \) only in \( h_1 \) and \( m_2|A \) intersects \( h \) only in \( h_2 \). Now, it is clear that

\[
m_1, m_2 \text{ intersect coherently in } A \text{ iff } m_1|A, m_2|A \text{ intersect coherently in } T_A.
\]

We abstract the present situation in the form of the following lemma characterizing coherent intersections of embedded curves in a torus relative to a fixed curve:

**Lemma 4.3.** Let \( T \) be a torus with oriented meridian-longitude curves \( \mu \) and \( \lambda \). Suppose that \( \lambda \) is the union of two arcs \( \lambda_1, \lambda_2 \), and that two embedded circles \( \gamma_1, \gamma_2 \) in \( T \) are given such that \( \gamma_i \) intersects \( \lambda \) in \( \text{int}(\lambda_i) \) only, for \( i = 1, 2 \). Subject only to this constraint, we further assume that \( \gamma_1 \) and \( \gamma_2 \) intersect minimally.

If \( \gamma_i = a_i\mu + p_i\lambda \) and \( a_i \geq 2 \), for \( i = 1, 2 \), then \( \gamma_1 \) and \( \gamma_2 \) intersect coherently iff \( \lceil p_1/a_1 \rceil \neq \lceil p_2/a_2 \rceil \). More precisely, we have that \( \gamma_1 \) and \( \gamma_2 \) intersect coherently with \( \gamma_1 \cdot \gamma_2 > 0 \) iff \( \lceil p_2/a_2 \rceil > \lceil p_1/a_1 \rceil \).

We remark that the conditions given in the lemma are independent of the meridian circle \( \mu \) used to specify coordinates in \( T \). The proof of the lemma is given at the end of this section, and we use it now to study the coherent intersection of the circles \( m_1, m_2 \) and \( m_1', m_2' \) in the annuli \( A \) and \( A' \), respectively. For such coherent intersections to represent all types, we must have that either:

(i) \( m_2 \) and \( m_2' \) intersect \( m_1 \) in the same direction and \( m_1' \) in opposite directions, or
(ii) \( m_2 \) and \( m'_2 \) intersect \( m'_1 \) in the same direction and \( m_1 \) in opposite directions.

We deal with (i) first, and refer to the frames of \( \partial V_1 \) and \( \partial V'_1 \) to compute all intersection numbers. From Figure 8 above, we can see that

\[
m_1 \cdot m'_2 > 0 \quad \text{and} \quad m'_1 \cdot m_2 > 0;
\]

therefore, (i) is equivalent to:

\[
(i') \quad m_1 \cdot m_2 > 0 \quad \text{and} \quad m'_1 \cdot m'_2 < 0.
\]

The coherent intersections of \( m_1, m_2 \) that satisfy (i') can be handled from the point of view of the framed torus \( T_A \) and the curves \( m_1|A, m_2|A \) via Lemma 4.3. To apply the lemma, though, the \( \mu_1 \)-coefficients of both curves must be positive to begin with. This is only a problem for \( m_2|A \), easily corrected by rephrasing (i') in terms of \(-m_2|A\); that is,

\[
\begin{align*}
m_1, m_2 \text{ intersect coherently} & \quad \text{iff} \quad m_1|A, m_2|A \text{ intersect coherently} \\
\quad \text{and} \quad m_1 \cdot m_2 > 0 & \quad \text{iff} \quad m_1|A \cdot (-m_2|A) < 0 \\
\text{iff} \quad \begin{bmatrix} -p_2/a_2 \end{bmatrix} < \begin{bmatrix} p_1/a_1 \end{bmatrix} \quad & \quad \text{iff} \quad c(M) = \begin{bmatrix} p_1/a_1 \end{bmatrix} + \begin{bmatrix} p_2/a_2 \end{bmatrix} > 0.
\end{align*}
\]

Similarly, \( m'_1, m'_2 \) intersect coherently and \( m'_1 \cdot m'_2 < 0 \) if \( c(M') < 0 \). Since it is now apparent that (ii) is equivalent to \( c(M) < 0 \) and \( c(M') > 0 \), we conclude that the manifold \( M \) gets a t-manifold structure from the given Heegaard splitting if \( C(M) = c(M)c(M') < 0 \). The Heegaard splitting (b) can be handled in a similar way and yields the same conclusion. The theorem follows. \( \square \)

Proof of Lemma 4.3. Let \( A \) be the annulus obtained by cutting \( T \) along \( \lambda \), and consider the arcs of \( \gamma_i \) in \( A \). Call any two such arcs \( \lambda_i \)-parallel if they cobound a band in \( A \) that is disjoint from the two copies of \( \lambda_j \) \( (j \neq i) \) in \( \partial A \). There can be simultaneously at most two disjoint arcs that are not \( \lambda_i \)-parallel. Since \( a_i \geq 2 \), it follows that the arcs of \( \gamma_i \) in \( A \) must split into two non \( \lambda_i \)-parallel families of \( \lambda_i \)-parallel arcs; amalgamate each of these families into one arc, and call the resulting two arcs \( x_i \) and \( y_i \). These two arcs give rise to a closed curve in \( T \) in the obvious way, which we denote by \( \overline{\gamma}_i \) and orient so as to follow the orientation of the arcs \( x_i, y_i \). The new curves \( \overline{\gamma}_i \)'s share the following properties in common with the \( \gamma_i \)'s:

- \( \gamma_1, \gamma_2 \) intersect coherently iff \( \overline{\gamma}_1, \overline{\gamma}_2 \) intersect coherently;
- in the case of coherent intersections, \( \gamma_1 \cdot \gamma_2 \) and \( \overline{\gamma}_1 \cdot \overline{\gamma}_2 \) have the same sign.

Both properties follow easily from the way the \( \overline{\gamma}_i \)'s were constructed. We now claim that:
Claim. $\gamma_1$ and $\gamma_2$ intersect coherently iff their intersection number in $T$ is nonzero.

Proof. To see this, let $x_i, y_i$ be the two arcs of $\gamma_i$ in $A$. We proceed according to the following two cases:

Case 1. No two arcs of $\gamma_1, \gamma_2$ in $A$ intersect in more than one point.

In this case, the situation must be as in Figure 10(a), where clearly the intersection number of the curves is zero and the intersection is not coherent.

Case 2. One arc of $\gamma_1$ intersects an arc of $\gamma_2$ in at least two points.

Suppose $|x_1 \cap x_2| \geq 2$; Figure 10(b) shows the annulus $A$ and all the arcs involved. Observe that the arc $x_2$ must intersect $x_1$ and $y_1$ in the same direction as $y_2$, inducing coherent intersections (and hence nontrivial intersection number). This proves our claim. □

To translate the information obtained about $\gamma_1$ and $\gamma_2$ in terms of $\gamma_1'$ and $\gamma_2'$, we proceed as follows. Choose the notation so that $\min\{|x_i \cap \mu|, |y_i \cap \mu|\} = |x_i \cap \mu| = n_i$. In order to construct the closed curve $\gamma_i$, the arc $y_i$ must be of the form $x_i \pm \lambda$, and so $|y_i \cap \mu| = |x_i \cap \mu| + 1 = n_i + 1$. Therefore,

$$\gamma_i = 2\mu + \text{sgn}(p_i)(2n_i + 1)\lambda.$$ 

Now suppose that $a_i'$ copies of $x_i$ and $a_i''$ copies of $y_i$ are needed to construct $\gamma_i$, where $a_i' + a_i'' = a_i$. Then,

$$\gamma_i = a_i\mu + p_i\lambda = (a_i' + a_i'')\mu + \text{sgn}(p_i)(n_i a_i' + (n_i + 1)a_i'')\lambda = a_i\mu + \text{sgn}(p_i)(n_i a_i' + a_i'')\lambda$$

Figure 10. The arcs $x_i$ and $y_i$ in the annulus $A$. 

(a)

(b)
whence \( p_i = \text{sgn}(p_i)(n_ia_i + a''_i) \), and so \( n_i = \lfloor p_i/a_i \rfloor \), where the brackets denote the greatest integer function. Observing that \( 2\lfloor x \rfloor = \text{sgn}(x)(2\lfloor |x| \rfloor + 1) \) for \( x \) not an integer, we can write

\[
\gamma_1 \cdot \gamma_2 \neq 0 \iff \text{sgn}(p_1)(2n_1 + 1) \neq \text{sgn}(p_2)(2n_2 + 1)
\]
and

\[
\gamma_1 \cdot \gamma_2 > 0 \iff \text{sgn}(p_2)(2n_2 + 1) - \text{sgn}(p_1)(2n_1 + 1) > 0
\]
iff \( \lfloor p_2/a_2 \rfloor > \lfloor p_1/a_1 \rfloor \).

The lemma follows. \( \square \)

5. Proper matrices.

In this section, we introduce some basic properties of proper matrices and give a proof of Theorem 1.3.

Two \( k \times n \) matrices are said to be \( t \)-equivalent if one can be obtained from the other by performing a finite sequence of operations of the following types:

(a) Permutating any row (column, resp.) with any other row (column, resp.);
(b) multiplying any row or column by \( \pm 1 \);
(c) substituting any nonzero entry by any nonzero real number of the same sign.

These operations will be referred to as \( t \)-equivalences of Types (a), (b), and (c). Observe that if \( A \) and \( A' \) are \( t \)-equivalent, then \( A \) is proper iff \( A' \) is proper; this fact will be used frequently. Given vectors \( v' \in \mathbb{R}^s \) and \( v'' \in \mathbb{R}^{n-s} \), we use the notation \( (v'|v'') \) to represent the vector in \( \mathbb{R}^n \) obtained by taking the entries of \( v' \) followed by those of \( v'' \).

Proof of Theorem 1.3 (cf. [9, Lemma 4.4]). Assume (c) does not hold, and let \( v \) be a nonzero vector in \( \mathbb{R}^n \) which is orthogonal to all the rows of \( A \). Let \( a \) be a row of \( A \) that represents \( v \). If all the entries of \( v \) are nonzero then \( a \) must be the zero vector, since \( \langle v, a \rangle = 0 \). As the zero vector represents all types, either \( A \) consists of a row of zeroes and satisfies (a), or else it must satisfy (b). Hence, passing to a \( t \)-equivalent matrix if necessary, we may assume that \( v \) is of the form \( (O|v'') \), where, for some \( 0 < s < n \), all the entries of \( v'' \in \mathbb{R}^{n-s} \) are nonzero, and \( O \in \mathbb{R}^s \) is the zero vector.

Let \( u' \) be any vector in \( \mathbb{R}^s \), and consider the vector \( u = (u'|v'') \). If \( a = (a'|a'') \) is any row of \( A \) that represents the vector \( u \), then \( a'' = O \) since \( \langle v, a \rangle = 0 \). Hence, the matrix \( B \) whose rows are the rows of \( A \) of the form \( (a'|O) \), \( O \in \mathbb{R}^{n-s} \), represents all \( n \)-types. If \( A = B \) then (a) holds, otherwise (b) holds. \( \square \)
In the context of proper matrices, Parry’s Theorem [13] can be restated as follows:

**Lemma 5.1 (Parry’s Theorem Restated).** Let $A$ be an integral $k \times n$ proper matrix. Then there is an $n \times n$ submatrix $A_0$ of $A$ such that $|\det(A_0)| \geq 2$.

**Proof.** By Parry’s Theorem, for some subset $S$ of the rows of $A$, the group $G_S = \mathbb{Z}^n/\langle S \rangle$ has nontrivial torsion. Let $S'$ be a maximal linearly independent subset of $S$; then $G'_{S'} = \mathbb{Z}^n/\langle S' \rangle$ also has nontrivial torsion. Since $A$ is proper, $\text{rank}(A) = n$ and so there is a set $T$ which consists of $n$ linearly independent rows of $A$ and contains $S'$. Since the group $G_T = \mathbb{Z}^n/\langle T \rangle$ necessarily has nontrivial torsion (and is finite), the matrix $A_0$ whose rows are the vectors in $T$ satisfies the conclusion of the lemma. \hfill $\square$

In the next lemma, we give a sufficient condition for a matrix to represent all types in terms of t-equivalence.

**Lemma 5.2.** Let $A$ be a $k \times n$ matrix. If any matrix that is t-equivalent to $A$ has rank $n$, then $A$ represents all types.

**Proof.** Suppose $A$ fails to represent some vector $v \in \mathbb{R}^n$; without loss of generality, and after a t-equivalence if necessary, we may assume that $v = (1, 1, \ldots, 1)$. Then every row of $A$ must have two entries with opposite signs. If $a$ is any row of $A$, let $a^*$ be a row vector obtained by substituting a pair of oppositely signed nonzero entries of $a$ with a pair of nonzero real numbers of the same signs, respectively, such that the sum of the entries of $a^*$ is zero, and let $A^*$ be the matrix whose rows are the $a^*$'s. Then clearly $A^*$ is t-equivalent to $A$, but $\text{rank}(A^*) < n$ since the column vectors of $A^*$ add up to zero. The lemma follows. \hfill $\square$

For any nonsingular $n \times n$ matrix $A = (a_{ij})$, we say that $\det(A)$ is coherent if all the nonzero terms in the usual expansion $\sum_{\sigma} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ of $\det(A)$ have the same sign.

**Lemma 5.3.** Let $A$ be an $n \times n$ nonsingular matrix. Then every matrix that is t-equivalent to $A$ is nonsingular iff $\det(A)$ is coherent.

**Proof.** Suppose $\det(A)$ is not coherent, say two terms $\text{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ and $\text{sgn}(\sigma') a_{1\sigma'(1)} \cdots a_{n\sigma'(n)}$ in the expansion of $\det(A)$ are nonzero and have opposite signs. Fix the numbers $a_{1\sigma(1)}, \ldots, a_{n\sigma(n)}$, and replace each of the other nonzero entries of $A$ by a sufficiently small number of the same sign. Obtain in this way a matrix $B$ that is t-equivalent to $A$, and such that $\det(B)$ and $\text{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ have the same sign. In a similar way, we can find a matrix $B'$ which is t-equivalent to $A$ and such that $\det(B')$ and $\text{sgn}(\sigma') a_{1\sigma'(1)} \cdots a_{n\sigma'(n)}$ have the same sign. Since the set $\mathcal{A}$ of all matrices which are t-equivalent to $A$ is connected (homeomorphic to $\mathbb{R}^k$, where $k$
is the number of nonzero entries of $A$) and the determinant function is continuous on $A$, it follows that $\det(A_0) = 0$ for some $A_0 \in A$.

Conversely, suppose that $A$ is nonsingular and $\det(A)$ is coherent, and let $A'$ be $t$-equivalent to $A$. Then there is a matrix $A''$ such that:

(i) $A''$ can be obtained via Type (c) $t$-equivalences on $A$, and

(ii) $A''$ can be obtained via Type (a) or (b) $t$-equivalences on $A'$.

It follows from (i) that $\det(A'')$ is coherent and hence that $A''$ is nonsingular. Since $|\det(A')| = |\det(A'')|$ by (ii), $A'$ is nonsingular as claimed. □

**Remark 5.4.** Implicit in Lemma 5.3 is the fact that, for nonsingular matrices, the property of having coherent determinant is invariant under $t$-equivalence.

Recall that a matrix is proper if it represents all types and does not satisfy (a) or (b) of Theorem 1.3. More generally, we say that a matrix $A$ represents types properly, and that it is $t$-proper for short, if each row of $A$ represents some type which is not represented by any other row of $A$. Clearly, proper matrices are $t$-proper, but the converse, in general, does not hold; observe however that a matrix is proper iff it is $t$-proper and represents all types. In this context, the previous two lemmas combine to prove the following result:

**Corollary 5.5.** Let $A$ be a nonsingular $t$-proper square matrix. Then $A$ is proper iff $\det(A)$ is coherent. □

The next result is a mild generalization of Parry's Theorem for square matrices.

**Lemma 5.6.** Let $A$ be a nonsingular $n \times n$ matrix ($n \geq 2$) with integral entries and coherent determinant. Then $|\det(A)| \geq 2$.

**Proof.** If $n = 2$ then $A$ is $t$-equivalent to a matrix of the form $(\pm \pm)$ by performing only Type (a) or (b) $t$-equivalences; hence $|\det(A)| \geq 2$. Proceeding by induction, suppose that $A = (a_{ij})$ is a nonsingular $n \times n$ matrix with $n > 2$ and coherent determinant. After performing Type (a) or (b) $t$-equivalences on $A$ if necessary, we may assume that both $a_{11}$ and its cofactor $C_{11}$ are nonzero. Let $M_{11}$ be the $(n-1) \times (n-1)$ minor corresponding to $a_{11}$, so that $\det(M_{11}) = C_{11}$. Since $\det(A) = a_{11}C_{11} + \ldots$ is coherent (see Remark 5.4), it follows that $\det(M_{11})$ is also coherent; as $M_{11}$ is nonsingular we get that $|\det(M_{11})| \geq 2$ by the induction hypothesis, and hence that $|\det(A)| \geq 2$ by coherency of $\det(A)$. □

A matrix that represent all types is proper iff it is $t$-proper. In the case of square matrices, it is possible to replace the later condition by one that involves only some knowledge about the $0$-girth of the matrix. This is the content of Theorem 1.5, whose proof is given below. We point out that the $0$-girth of a square matrix is invariant under $t$-equivalence, a fact which is used implicitly in the proof.
**Proof of Theorem 1.5.** Suppose $A$ represents all types but is not proper; then, by definition, $A$ must either have a column of zeroes or some proper subset of its rows must represent all types. In the first case the 0-girth of $A$ is at least $n + 1$ (the girth of any zero-column); in the second case, let $A_0$ be a $k \times n$ submatrix of $A$ that represents all types, with $k < n$ smallest among all choices. After a $t$-equivalence, if necessary, we may assume that $A_0 = (B \mid O)$, where $B$ is a $k \times l$ matrix which represents all types and has no zero columns. The minimality of $k$ then implies that $B$ is proper and hence that $l \leq k$ by Theorem 1.3. Since $A$ is $t$-equivalent to a matrix of the form

$$
(B \quad O) \quad O = \text{zero matrix},
$$

the 0-girth of $A$ is at least $k + (n - l) \geq n$. Hence, in all cases, the 0-girth of $A$ is greater than $n - 1$.

For the converse, assume that $A$ is proper. If the 0-girth of $A$ is at least $n$ then, after performing a $t$-equivalence if necessary, $A$ may be assumed to be of the form given in (5.1), where $B$ is a $k \times l$ matrix with $l \leq k < n$; observe that $k = l > 0$ since $A$ is nonsingular, so $B$ and $D$ are square matrices. Since $A$ is proper, each matrix $t$-equivalent to $A$ is proper and hence nonsingular. From the equation $\det(A) = \det(B) \det(D)$ we can then see that $B$ and every matrix $t$-equivalent to $B$ must also be nonsingular. Hence, by Lemma 5.2, $B$ represents all types, contradicting the properness of $A$. □

The $n \times n$ matrix

$$
A_n = \begin{pmatrix}
+ & + & \ldots & \ldots & + & + \\
+ & + & \ldots & \ldots & + & - \\
+ & + & \ldots & \ldots & - & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
+ & + & - & 0 & \ldots & 0 \\
+ & - & 0 & \ldots & 0
\end{pmatrix},
$$

where the signs of the nonzero entries are as shown, is proper and its 0-girth is $n - 1$, which shows that the upper bound for the 0-girth of a proper matrix given in Theorem 1.5 is the best possible.

Combining Theorem 1.5 with Lemmas 5.2 and 5.3, we get the following characterization of proper square matrices:

**Corollary 5.7.** Let $A$ be an $n \times n$ nonsingular matrix, $n \geq 2$. Then $A$ is proper iff $\det(A)$ is coherent and the 0-girth of $A$ is at most $n - 1$. □

**Corollary 5.8.** Let $A$ be a square matrix. Then $A$ is proper iff $A^T$ is proper.

The family of proper square matrices $A_n$ given above is part of the bigger family of proper square matrices that have $2 \times 2$ submatrices not representing
all types. We will see shortly that every member of this bigger family also has maximal 0-girth. The first step in this direction consists in showing that singular square matrices which remain singular under t-equivalences must have ‘large’ 0-girths.

**Lemma 5.9.** Let $A$ be an $n \times n$ matrix. Then the 0-girth of $A$ is at least $n + 1$ iff every matrix that is t-equivalent to $A$ is singular.

This is a well-known result in combinatorics which goes back to P. Hall’s Theorem [10] and the *marriage problem*; a short proof of Hall’s Theorem can be found in the paper of Halmos and Vaughn [11], and a proof of Lemma 5.9 can be found in the book of C. Berge [1, Theorem 9, p. 105].

We now restrict our attention to the family of proper square matrices that have $2 \times 2$ submatrices not representing all types.

**Lemma 5.10.** Let $A$ be an $n \times n$ proper matrix with $n > 2$. If some $2 \times 2$ submatrix of $A$ does not represent all types, then the 0-girth of $A$ is $n - 1$.

**Proof.** After a t-equivalence, if necessary, $A$ may be assumed to be of the form
\[
\begin{pmatrix}
 a & b & \cdots \\
 c & d & \cdots \\
 \vdots & \vdots & \ddots \\
 & & & B
\end{pmatrix},
\]
where $a, b, c, d$ are all positive real numbers and $B$ is an $(n - 2) \times (n - 2)$ matrix. Then $\det(A) = (ad - bc) \det(B) + k_1a + k_2b + k_3c + k_4d + k$, where the numbers $\det(B), k_1, k_2, k_3, k_4$, and $k$ do not depend on $a, b, c,$ or $d$.

After some some t-equivalence of $A$, if necessary, we may assume that $\det(B) \geq 0$. For any positive real number $x$, let $A_{x,+}$ and $A_{x,-}$ be the matrices t-equivalent to $A$ obtained by assigning the values $a = x, b = 1, c = 1, d = x$ and $a = 1, b = x, c = x, d = 1$, respectively. If $\det(B) > 0$ then
\[
\lim_{x \to \infty} \det(A_{x,+}) = \infty \quad \text{and} \quad \lim_{x \to -\infty} \det(A_{x,-}) = -\infty,
\]
so that $\det(A)$ must be zero at some ‘point’ of the connected set $\mathbb{R}_+^4 = \{(a, b, c, d) \mid a, b, c, d > 0\}$, which, in light of Theorem 1.3 and the invariance of properness under t-equivalences, contradicts the properness of $A$. Hence $\det(B) = 0$, and similarly $\det(B') = 0$ whenever $B'$ is t-equivalent to $B$, which by Lemma 5.9 implies that the 0-girth of the $(n - 2) \times (n - 2)$ matrix $B$, and hence of $A$, must be at least $n - 1$. That $A$ has 0-girth exactly equal to $n - 1$ now follows from Theorem 1.5. \[\square\]

A word about standard notation in combinatorial matrix theory is appropriate now (cf. [2]). An $n \times n$ matrix $A$ is *sign-nonsingular* if $\det(A)$ is coherent (equivalently, if $|\per(A)| = |\det(A)|$, where $\per(A)$ denotes the *permanent* of $A$); if $A$ is a $(0, 1)$-matrix, then $A$ is *indecomposable* if its 0-girth is at most $n - 1$. Thus, any nonsingular $(0, 1)$-matrix is proper iff it is
Figure 11. Lines in $X = P^2(\mathbb{F}_2)$.

sign-nonsingular and indecomposable (Corollary 5.7), and any square matrix
that represents all types is proper iff it is indecomposable (Theorem 1.5).

A finite projective plane of order $q \geq 1$ is an $n \times n$ $(0, 1)$-matrix $A$ that
satisfies the equation $AA^T = qI + J$, where $n = q^2 + q + 1$ and $J$ is the
matrix all of whose entries are 1’s (cf. [2, §1.3]). The smallest projective
plane corresponds to the value $q = 1$ and is given by the proper matrix

$$
\begin{pmatrix}
1 & 1 & 0 \\
1 & 0 & 1 \\
0 & 1 & 1
\end{pmatrix}.
$$

When $q$ is the power of a prime, an example of such a matrix can be con-
structed using the points and lines of the projective plane $P^2(\mathbb{F}_q)$, where $\mathbb{F}_q$
is the field of $q$ elements. For instance, in the case $q = 2$, there is only one
finite projective plane (up to permutation of rows or columns), obtained as
follows:

Let $X = P^2(\mathbb{F}_2)$ denote the projective space of the 3-dimensional vector
space $\mathbb{F}_2^3$ over the field $\mathbb{F}_2$ of two elements. The space $X$ consists of 7 points
which are organized into subsets that represent the lines of $X$ (the images of
2-dimensional subspaces of $\mathbb{F}_2^3$). The set $X$ is shown in Figure 11, where each
point of $X$ has been identified with an integer from 1 to 7; any three points
connected by a line or circle form a line in $X$. For each line $l = \{i, j, k\}$,
$i < j < k$, of $X$, let $[l]$ be the ordered 7-tuple consisting of 0’s and +’s,
whose +-entries occur at positions $i, j, k$, and let $A_X$ be the matrix whose
rows are the vectors $[l]$ obtained in this way. The matrix $A_X$ is called the
incidence matrix of \( X \). Hence, up to \( t \)-equivalence,

\[
A_X = \begin{pmatrix}
0 & + & + & 0 & + & 0 & 0 \\
+ & 0 & + & 0 & 0 & + & 0 \\
+ & + & 0 & 0 & 0 & 0 & + \\
0 & 0 & 0 & 0 & + & + & + \\
+ & 0 & 0 & + & + & 0 & 0 \\
0 & + & 0 & + & + & 0 & + \\
0 & 0 & + & + & 0 & 0 & +
\end{pmatrix}.
\]

We can see that \( A_X \) is proper as follows: Let \( \varepsilon \) be any 7-type, and color red and green its positive and negative entries, respectively. Then \( A_X \) represents \( \varepsilon \) if a set of three like-colored entries of \( \varepsilon \) correspond with the locations of \( + \)'s in some row of \( A_X \). Suppose \( \varepsilon \) has three red-colored entries which do not correspond to any \( + \)'s in any row of \( A_X \), say they correspond to the subset \( 1, 2, 3 \) of \( X \) (which is not a line). Then if any of the entries of \( \varepsilon \) in positions 5, 6, or 7 is colored red we are done, for the sets \( 1, 2, 7, 2, 3, 5 \), and \( 1, 3, 6 \) are lines of \( X \). Otherwise, all three positions 5, 6, and 7 of \( \varepsilon \) must be colored green and again we are done. By homogeneity of the space \( X \), the same argument must work in all other cases, which proves that \( A_X \) represents all types. That \( A_X \) is proper now follows from Theorem 1.5 since the 0-girth of \( A_X \) is 5; in particular, the results in this section imply that \( A_X \) is sign-nonsingular, a property deemed ‘unusual’ in \([2, \text{s1.3}]\).

Higher order finite projective planes do not share the properness nor the sign-nonsingularity of the small cases \( q \leq 2 \); we leave the details of this fact to the interested reader, in the form of an exercise:

**Exercise 5.11.** Let \( A \) be a finite projective plane of order \( q \geq 1 \). Then:

(a) \( A \) is indecomposable; in fact, the 0-girth of \( A \) is \( q^2 + 1 \).

(b) \( A \) represents all types (equivalently, \( A \) is sign-nonsingular) iff \( q \leq 2 \).

6. **Topology of \( t \)-manifolds.**

We now apply the results on proper matrices and representation of types established in the previous section to the study of general \( t \)-manifolds. Let \( M \) denote a \( t \)-manifold with a genus \( n \) \( t \)-manifold structure \( S = (H_n, D, C) \) and associated integral matrix \( A_M \). Recall that \( A_M \) is a presentation matrix for the group \( H_1(M) \); any presentation matrix of \( H_1(M) \) obtained in this way will be called a \( t \)-presentation matrix of \( M \). Recall that \( A_M \) is proper.

**Proof of Theorem 1.4** (cf. \([6, \text{Theorem 4.3}]\)). The result is obvious for genus one \( t \)-manifold structures with the only exception of \( S^3 \), which the hypothesis \( A_M \neq (1) \) excludes.

Suppose \( M \) has a genus \( n \) \( t \)-manifold structure, where \( n > 1 \); then (a) follows directly from the properness of \( A_M \), hence \( M \) is closed and \( H_1(M) \)
is finite. Since any rank $n$ submatrix of $A_M$ is a presentation matrix for $H_1(M)$, that $H_1(M)$ is nontrivial follows from Lemma 5.1.

Let $\{H, D, C\}$ be a t-manifold structure of some t-manifold $M$. Since, by Theorem 1.4, the manifold $M$ is closed, the set $H' = M \setminus H$ is a handlebody and hence the pair $(H, H')$ is a Heegaard splitting of $M$. That the curves in $C$ intersect the meridians of $H$ coherently easily proves the following result:

**Corollary 6.1.** Let $M$ be a t-manifold with a $k \times n$ t-presentation matrix $A_M$. If any of the entries of $A_M$ is $\pm 1$, then the Heegaard genus of $M$ is strictly less than $n$. □

Some more information on the topology of a given t-manifold can be obtained if a suitable t-manifold structure is at hand. For example, many of the t-manifold structures in Section 3 arise from Heegaard diagrams of minimal genus. We will see how to partially generalize this situation to higher genus t-manifold structures in the cases when the t-presentation matrix is square, where some of the results of Section 5 on 0-girths of proper matrices can be exploited to give topological information on the t-manifold.

Let $M$ be a t-manifold with t-manifold structure $S = \{H_n, D, C\}$ and t-presentation matrix $A_M$. Then $C$ has at most $3n - 3$ components and represents all $n$-types. Now, the $2^n$ different $n$-types come in pairs $\pm \epsilon$, and clearly every $k \times n$ proper matrix with no zero entries must be t-equivalent to the matrix whose rows are the vectors $\pm \epsilon$ just described; in particular, $k = 2^{n-1}$. We combine these two simple observations with the main theorem of [3] to prove the following result:

**Lemma 6.2.** Let $M$ be a t-manifold with a genus $n$ t-manifold structure $\{H_n, D, C\}$ and associated Heegaard splitting $(H_n, H'_n)$. If $n \geq 5$ then the splitting $(H_n, H'_n)$ is not strongly irreducible. In particular, either $(H_n, H'_n)$ is reducible or $M$ contains an incompressible closed surface of positive genus.

**Proof.** Since $3n - 3 < 2^{n-1}$ for $n \geq 5$, our previous observations imply that $A_M$ must have at least one zero entry. This implies that for some circle $c \in C$ and disk $D \in D$, $c \cap D = \emptyset$, hence the splitting is not strongly irreducible and the last part of the lemma follows from [3, Theorem 3.1]. □

The conclusion of the previous lemma can be considerably strengthened in the special case when the t-manifold structure gives a Heegaard splitting of minimal genus. This is the content Theorem 1.6, whose proof we proceed to give. In the process, we will follow the notation of [3, §3] closely.

**Proof of Theorem 1.6.** Let the t-manifold structure of $M$ be given by the data $S = \{H_n, D = \{D_1, \ldots, D_n\}, C = \{c_1, \ldots, c_n\}\}$ and let $H'_n$ be the handlebody in $M$ complementary to $H_n$. We think of the circles in $C$ as the meridian circles of $H'_n$, which bound disjoint properly embedded disks
\(D' = \{D'_1, \ldots, D'_n\}\) in \(H'_{n}\). Since the Heegaard genus of \(M\) is \(n\), \(M\) will be reducible once the Heegaard splitting \((H_n, H'_{n})\) is reducible.

Recall from Theorem 1.5 that the 0-girth of \(A_M\) is at most \(n - 1\). If the 0-girth of \(A_M\) is \(n - 1\), then there are \(k\) meridian disks of \(H_n\), say \(D^* = \{D_1, \ldots, \partial D_k\}\), and \(l\) meridian disks of \(H'_{n}\), say \(C^* = \{C'_1, \ldots, C'_l\}\), which are pairwise disjoint, such that \(k, l > 0\) and \(k + l = n - 1\). Observe that the circles \(\partial C^* \cup \partial D^*\) are homologically independent in the Heegaard surface \(F = \partial H_n\); hence \(T = \sigma(F; \partial C^* \cup \partial D^*)\) is a torus and so \(c(\partial C^* \cup \partial D^*) = 2n - 2\).

If the torus \(T\) is incompressible in \(M\) we are done, so we may assume \(T\) compresses. By the proof of [3, Theorem 3.1], it follows that either the Heegaard splitting \((H_n, H'_{n})\) is reducible or there are nonempty collections \(E, E'\) of disjoint, properly embedded disks in \(H_n, H'_{n}\), respectively, such that

\[
\partial E \cap \partial E' = \emptyset \quad \text{and} \quad c(\partial E \cup \partial E') > c(\partial E), c(\partial E'),
\]

with \(c(\partial E \cup \partial E') > c(\partial C^* \cup \partial D^*) = 2n - 2\). But, for any 1-submanifold \(\alpha\) of \(F\), \(c(\alpha) \leq 2n - 1\), with equality holding only if \(\sigma(F; \alpha)\) consists of 2-spheres. Hence \(\sigma(F; \partial E \cup \partial E')\) is a union of 2-spheres and, by the proof of [3, Theorem 3.1], the Heegaard splitting \((H_n, H'_{n})\) again must be reducible.

The last part of the theorem follows immediately from Lemma 5.10. \(\square\)

**Proof of Theorem 1.7.** We assume here that \(X(s)\) has no lens space connected summand; the case when \(X(s)\) has no t-manifold summand is similar (cf. [15, Theorem 1.3]). By [6, Theorem 4.3], \(X(r)\) always has a t-manifold summand.

Suppose that \(X(r)\) has two prime connected summands \(M_1\) and \(M_2\), neither of which is part of a t-manifold summand. Let \(\tilde{P}\) be a 2-sphere in \(X(r)\) which ‘separates’ \(M_1\) from \(M_2\), and such that \(P = \tilde{P} \cap X\) is properly embedded in \(X\) with \(|\partial P|\) smallest subject to these constraints; then \(P\) is essential in \(X\) with boundary slope \(r\). Let \(Q\) be an essential planar surface in \(X\) with boundary slope \(s\); we may assume that \(P\) and \(Q\) intersect transversely and that any circle component of \(P \cap Q\) is nontrivial in both \(P\) and \(Q\). Since \(X(s)\) has no lens space connected summand, the graph \(G_Q = P \cap Q \subset Q\) represents all types by [8]. Let \(\Sigma\) be a generalized Scharlemann cycle of \(G_Q\); attaching \(\Sigma\) to \(\tilde{P}\) (see the Introduction) produces a 2-sphere \(\tilde{P}'\) in \(X(r)\) which cobounds a t-manifold with \(\tilde{P}\) and satisfies \(|\partial P'| < |\partial P|\). Since neither manifold \(M_1\) nor \(M_2\) is part of a t-manifold summand of \(X(r)\), it follows that \(\tilde{P}'\) ‘separates’ \(M_1\) from \(M_2\) in \(X(r)\), contradicting the minimality of \(|\partial P|\). The theorem follows. \(\square\)

We end this section with two examples of proper matrices and a couple of open questions.
Consider the matrix
\[
A = \begin{pmatrix}
1 & 1 & 1 \\
1 & 1 & -1 \\
1 & -1 & 1 \\
1 & -1 & -1
\end{pmatrix}.
\]

Figure 12 shows two t-manifold structures with t-presentation matrix \(A\); in the case of Figure 12(a) and (b) the t-manifold has fundamental group isomorphic to \(\mathbb{Z}/2\mathbb{Z} \ast \mathbb{Z}/2\mathbb{Z}\) and \(Q_8\), respectively, where \(Q_8 = \langle u, v \mid u^2 = (uv)^2 = v^2 \rangle\) is the classical quaternion group; hence the manifold in Figure 12(a) is homeomorphic to \(P^3 \# P^3\) (see Corollary 6.1). In particular, there exist reducible t-manifolds.

It can be proved that the fundamental group of any t-manifold with t-presentation matrix \(A\) is isomorphic to one of the two groups above [14].

Another interesting example involves the matrix
\[
F = \begin{pmatrix}
1 & 1 & 1 & 0 \\
1 & -1 & 0 & 1 \\
1 & 0 & -1 & -1 \\
0 & 1 & -1 & 1
\end{pmatrix}.
\]

It is not hard to see that any proper 4×4 matrix with 0-girth at most 2 is t-equivalent to \(F\). However, \(F\) is not a t-presentation matrix for any t-manifold \(M\) with t-manifold structure \(S = \{H_4, D = \{D_1, \ldots, D_4\}, C = \{c_1, \ldots, c_4\}\}\). To see this, let \(x_1, \ldots, x_4\) represent generators of \(\pi_1(H_4)\) dual to the \(D_i\)'s, respectively; then \(\pi_1(M) = \langle x_1, \ldots, x_4 \mid c_1, \ldots, c_4 \rangle\). Due to the
coherency of the intersections between the $\partial D_i$’s and the $c_j$’s, there are only a few choices for each word $c_i$. For example, we may have that, up to cyclic permutation,

$$
c_1 = xyz, \ c_2 = xy^{-1}w, \ c_3 = xz^{-1}w^{-1}, \ c_4 = yz^{-1}w,$$

which would yield

$$
\pi_1(\mathcal{M}) \cong G = \langle x, t \mid x^3t^3 = 1, \ txt^{-1} = x^{-2} \rangle
$$

after substituting $w = xt$. A little more computation shows that all the possible groups $\pi_1(\mathcal{M})$ obtained in this way are isomorphic to $G [14]$; hence in all cases $\pi_1(\mathcal{M})$ is nilpotent of order 27, and $H_1(\mathcal{M}) \cong \mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$. This contradicts the fact that any finite nilpotent 3-manifold group is cyclic or has even order [5], which proves our claim.

The above example of the $4 \times 3$ proper matrix $A$ shows that a t-manifold may be reducible; in that particular case, however, we can see that the prime factors of the t-manifold are also t-manifolds (in fact, lens spaces). So we ask:

**Question 1.** Are the prime factors of any t-manifold also t-manifolds?

A positive answer to this question would improve the conclusion of Theorem 1.7 to say: _If $X(s)$ has no lens space connected summands, then at most one of the prime factors of $X(r)$ is not a t-manifold._

By Theorem 1.1, no t-manifold with a genus two t-manifold structure is hyperbolic, and Theorem 1.6 provides many more examples of non-hyperbolic t-manifolds. So we ask:

**Question 2.** Are there any hyperbolic t-manifolds?

A negative answer to this last question would imply that t-manifolds are exceptionally rare, as would be the reducible surgeries giving rise to them.

References


Received June 1, 1999.

DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF TEXAS AT EL PASO
EL PASO, TX 79968
E-mail address: valdez@math.utep.edu
CONTENTS

Volume 210, no. 1 and no. 2

Konstantin Athanassopoulos: Explosions near isolated unstable attractors 201

E.N. Dancer, Yihong Du and Li Ma: Asymptotic behavior of positive solutions of some elliptic problems 215

Michael Dorff, Denise Halverson and Gary Lawlor: Area-minimizing minimal graphs over nonconvex domains 229

Yihong Du with E.N. Dancer and Li Ma 215

Kirsten Eisenträger: Hilbert’s tenth problem for algebraic function fields of characteristic 2 261

F.T. Farrell and L.E. Jones: Local collapsing theory 1

Roberto Frigerio, Bruno Martelli and Carlo Petronio: Complexity and Heegaard genus of an infinite class of compact 3-manifolds 283

Shamindra Kumar Ghosh: Higher exchange relations 299

Denise Halverson with Michael Dorff and Gary Lawlor 229

Kazuhiro Ichihara and Masakazu Teragaito: Klein bottle surgery and genera of knots 317

L.E. Jones with F.T. Farrell 1

Marc Lackenby: Exceptional surgery curves in triangulated 3-manifolds 101

Gary Lawlor with Michael Dorff and Denise Halverson 229

Li Ma with E.N. Dancer and Yihong Du 215

Bruno Martelli with Roberto Frigerio and Carlo Petronio 283

J.-A. Marti: Nonlinear algebraic analysis of delta shock wave solutions to Burgers’ equation 165

J. Mc Laughlin: Multi-variable polynomial solutions to Pell’s equation and fundamental units in real quadratic fields 335

Carlo Petronio with Roberto Frigerio and Bruno Martelli 283

Luis Gerardo Valdez Sánchez: Representation of types and 3-manifolds 367

Ichiro Shimada: The fundamental group of the complement of a resultant hypersurface 351
A. Stoimenow (with an Appendix by Mark Kidwell): *The crossing number and maximal bridge length of a knot diagram* 189

Masakazu Teragaito with Kazuhiro Ichihara 317

Ichiro Torisu: *On the additivity of the Thurston–Bennequin invariant of Legendrian knots* 359
Guidelines for Authors

Authors may submit manuscripts at pjm.math.berkeley.edu/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LATEX, but papers in other varieties of TEX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LATEX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTEX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this:”). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal’s preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.
Explosions near isolated unstable attractors 201
KONSTANTIN ATHANASSOPOULOS

Asymptotic behavior of positive solutions of some elliptic problems 215
E. N. DANCER, YIHONG DU AND LI MA

Area-minimizing minimal graphs over nonconvex domains 229
MICHAEL DORFF, DENISE HALVORSON AND GARY LAWLER

Hilbert’s tenth problem for algebraic function fields of characteristic 2 261
KIRSTEN EISENTRÄGER

Area-minimizing minimal graphs over nonconvex domains 229
MICHAEL DORFF, DENISE HALVORSON AND GARY LAWLER

Hilbert’s tenth problem for algebraic function fields of characteristic 2 261
KIRSTEN EISENTRÄGER

Multi-variable polynomial solutions to Pell’s equation and fundamental units 335
J. M. LAUGHLIN

The fundamental group of the complement of a resultant hypersurface 351
ICHIRO SHIMADA

On the additivity of the Thurston–Bennequin invariant of Legendrian knots 359
ICHIRO TOSHI

Representation of types and 3-manifolds 367
LUIS GERARDO VALDEZ SÁNCHEZ