**HAMBURGER BEITRÄGE**
**ZUR MATHEMATIK**

Heft 251

# MacLane's Theorem for Higher Surfaces

H. Bruhn, Grenoble
R. Diestel, Hamburg

Juni 2006

# MacLane's theorem for higher surfaces

Henning Bruhn* and Reinhard Diestel

**Abstract**

We generalise MacLane's planarity criterion to surfaces of higher genus.

## 1 Introduction

MacLane's well-known planarity criterion characterises the finite planar graphs in terms of their cycle space. As the *cycle space* $\mathcal{C}(G)$ of a graph $G$ we take the $\mathbb{Z}_2$-vector space generated by the edge sets of cycles in $G$, with symmetric difference as addition. Its elements are those sets $F \subseteq E(G)$ such that every vertex of $G$ is incident with an even number of edges in $F$. Call a family $\mathcal{F}$ of sets $F \subseteq E(G)$ *sparse* if every edge of $G$ lies in at most two members of $\mathcal{F}$.

MacLane's planarity criterion can then be stated as follows:

**Theorem 1 (MacLane [7]).** *A finite graph is planar if and only if its cycle space is generated by some sparse family of (edge sets of) cycles.*

The aim of this paper is to generalise MacLane's theorem to arbitrary closed surfaces.

Our approach is motivated by simplicial homology, as follows. Let a graph $G$ be embedded in a closed surface $S$ of minimal Euler genus $\varepsilon := 2-\chi(S)$. Then $S$ can be viewed as the underlying space of a 2-dimensional CW-complex $C$ with 1-skeleton $G$. Its first homology group $Z_1(C;\mathbb{Z}_2)/B_1(C;\mathbb{Z}_2)$ is $\mathbb{Z}_2^\varepsilon$, the direct product of $\varepsilon$ copies of $\mathbb{Z}_2$.

In graph theoretic language this means that the subspace $\mathcal{B}$ $(= B_1(C;\mathbb{Z}_2))$ spanned in $\mathcal{C}(G)$ $(= Z_1(C;\mathbb{Z}_2))$ by the set of face boundaries of $G$ in $S$ has codimension $\varepsilon$ in $\mathcal{C}(G)$. Now the set of face boundaries is a sparse set of cycles. Thus, if $G$ embeds in a surface of small Euler genus, at most $\varepsilon$, then $G$ has a sparse set of cycles spanning a large subspace in $\mathcal{C}(G)$, one of codimension at most $\varepsilon$.

MacLane's theorem says that, for $\varepsilon = 0$, the converse implication holds too: if $G$ has a sparse set of cycles whose span in $\mathcal{C}(G)$ has codimension at most $\varepsilon = 0$, then $G$ embeds in the (unique) surface of Euler genus at most $\varepsilon = 0$, the sphere. Our initial aim, then, would be to prove this converse implication for arbitrary $\varepsilon$.

This naive extension soon runs into difficulties, and indeed is not true. In Section 3 we discuss the obstructions encountered as they arise, and modify our naive conjecture accordingly. The result will be a collection of theorems, presented in Section 4, which each characterise embeddability in a surface of

*Supported by the Algorithmic Discrete Optimization Network (ADONET).

given Euler genus by a condition akin to MacLane's planarity criterion that is both necessary and sufficient. All proofs are given in Section 5.

Some previous work in this direction can be found in the literature. Mohar [8] also starts out from the above-mentioned necessary condition for embeddability in a surface of Euler genus at most $\varepsilon$, that the graph must have a sparse set of cycles whose span in its cycle space has codimension at most $\varepsilon$. Unlike our plan here, Mohar does not strengthen this condition to one that is also sufficient, but establishes how much it implies as it is; the (best possible) result is that it implies embeddability in a surface of Euler genus at most $2\varepsilon$. Edmonds [5] also studies embeddability in higher surfaces, in terms of duality. Our results will imply Edmonds', but we defer this to a separate paper on extensions of Whitney's theorem to higher surfaces [3].

# 2 General definitions and background

All graphs we consider are finite. Our notation will be that of [4], except that instead of 'multigraph' we say 'graph'. (Thus, our graphs may have loops and multiple edges, and degrees and connectivity are defined as they are in [4] for multigraphs. In particular, 2-connected graphs cannot have loops.) In the statements of some of our results we do not allow loops, but only to avoid unnecessary complication in our terminlogy: those theorems can be applied to graphs with loops by subdividing (and thereby eliminating) these.

The set of edges of a graph $G = (V, E)$ incident with a given vertex $v$ is denoted by $E(v)$. When $W$ is a walk in $G$, we denote the subgraph of $G$ that consists of the edges on $W$ and their incident vertices by $G[W]$; note that this need not be an induced subgraph of $G$. Both the edge space of $G$ and its subspace, the cycle space $\mathcal{C}(G)$, are taken over $\mathbb{Z}_2$. We write their elements as subsets of $E$, so vector addition becomes symmetric difference of edge sets. If $G$ is connected and has $n$ vertices and $m$ edges, its cycle space has dimension

$$\dim \mathcal{C}(G) = m - n + 1. \tag{1}$$

A (closed) *surface* is a compact connected 2-manifold without boundary.

An *$n$-dimensional CW-complex*, or *$n$-complex*, is a finite set $C$ of open balls $B^i_j \subseteq \mathbb{R}^i$ with $i \le n$, called *$i$-cells*, that have disjoint closures and whose union is made into a topological space $|C|$ as follows. The union $C^0$ of all 0-cells (which are singletons, so $C^0$ is just a set of points) carries the discrete topology. Assume now that the union of all $i$-cells with $i \le k < n$, the *$k$-skeleton* $C^k$ of $C$, has been given a topology, and denote this space by $|C^k|$. For every $(k+1)$-cell $B^{k+1}_j \in C$ choose a continuous *attachment map* $f_j : \partial B^{k+1}_j \to |C^k|$ from its boundary $\partial B^{k+1}_j$ in $\mathbb{R}^{k+1}$ to $|C^k|$. Then give $|C^{k+1}|$ the quotient topology of the (disjoint) union of $|C^k|$ with all the closures of the $B^{k+1}_j$ obtained by identifying every $x \in \partial B^{k+1}_j$ with $f_j(x)$.

Every graph $G$ is a 1-complex, with vertices as 0-cells and edges as 1-cells. A topological embedding of $G$ in another space $S$ is a *2-cell-embedding* if $G$ is the 1-skeleton of a 2-complex $C$ such that the embedding of $G$ in $S$ extends to a homeomorphism $\varphi\colon |C| \to S$. The images under $\varphi$ of the 2-cells of $C$ are the *faces* of $G$ in $S$. If $S$ is a surface, their attachment maps define closed walks in $G$. These walks are unique up to cyclic shifts and orientation, a difference we shall ignore. We thus have one such walk assigned to each face, and call this family the (unique) *family of facial walks*. If $W$ is the facial walk of some face $f$, then $\varphi$ maps the subgraph $G[W]$ onto the frontier of $f$ in $S$, and we call $G[W]$ the *boundary* of the face $f$.

Given a surface $S$, consider any 2-cell-embedding of any graph in $S$. Let $n$ be its number of vertices, $m$ its number of edges, and $\ell$ its number of faces in $S$. Euler's theorem tells us that $n - m + \ell$ is equal to a constant $\chi(S)$ depending only on $S$ (not on the graph), the Euler characteristic of $S$. The *Euler genus* $\varepsilon(S)$ of $S$ is defined as the number $2 - \chi(S)$. Euler's theorem then takes the following form, which we refer to as *Euler's formula*:

$$\varepsilon(S) = m - n - \ell + 2. \tag{2}$$

Given a graph $G$, let $\varepsilon = \varepsilon(G)$ be minimal such that $G$ has a topological embedding $\varphi$ in a surface of Euler genus at most $\varepsilon$. This $\varepsilon$ is the *Euler genus of* $G$, and any such $\varphi$ is a *genus-embedding* of $G$. Genus-embeddings of connected graphs are 2-cell-embeddings [9, p. 95]. If $G$ has components $G_1, \dots, G_n$, then $\varepsilon(G) = \varepsilon(G_1) + \cdots + \varepsilon(G_n)$, a fact referred to as *genus additivity* [9]. (The same is true for blocks rather than components, but we do not need this.)

We say that a family $\mathcal{W}$ of walks *covers* $G$ if every edge of $G$ lies on some walk of $\mathcal{W}$. It covers an edge $e$ *$k$ times* if $k = \sum_{W \in \mathcal{W}} k_W(e)$, where $k_W(e)$ is the number of occurrences of $e$ on $W$ (irrespective of the direction in which $W$ traverses $e$). $\mathcal{W}$ is a *double cover* of $G$ if it covers every edge of $G$ twice. A walk is *non-trivial* if it contains an edge.

Given a walk $W$ in $G$, we write $c(W)$ for the set of edges that appear an odd number of times in $W$. Note that, if $W$ is closed, then $c(W)$ lies in $\mathcal{C}(G)$, the cycle space of $G$. The *dimension* of a family $\mathcal{W}$ of closed walks, $\dim \mathcal{W}$, is the dimension of the subspace spanned in $\mathcal{C}(G)$ by the sets $c(W)$ with $W \in \mathcal{W}$. The *codimension* of $\mathcal{W}$ in $\mathcal{C}(G)$ is the number $\dim \mathcal{C}(G) - \dim \mathcal{W}$.

## 3 Reconstructing a surface

MacLane's theorem offers a necessary and sufficient condition for embeddability in a fixed surface, the sphere. Ideally, we would like to have a similar condition characterising embeddability in an arbitrary but fixed surface $S$.

To illustrate what we mean by 'similar', let us think of MacLane's theorem as listing some properties of the facial cycles of a plane graph—sparseness and generating the entire cycle space—which, together, are strong enough to imply the following: that whenever we have *any* collection of cycles with these properties and attach a 2-cell to each of them, the 2-complex obtained is homeomorphic to the sphere. (This, indeed, is the outline of the standard topological proof of MacLane's theorem.)

For an arbitrary surface $S$, we are thus looking for a similar list of properties shared by the facial cycles of all graphs suitably embedded in $S$ (with a genus-embedding, say) that allows us to reconstruct $S$ by attaching a 2-cell along each of those cycles. One of those properties should be sparseness: if more than two 2-cells meet in an edge, the complex obtained will not be a surface. Following the homological approach outlined in the introduction, we might complement this by requiring that our cycles span a large enough subspace of the cycle space:

**Naive Conjecture.** *A graph $G$ embeds in a surface $S$ if and only if $G$ has a sparse set of cycles whose span in $\mathcal{C}(G)$ has codimension at most $\varepsilon(S)$ in $\mathcal{C}(G)$.*

Notice that this conjecture can be true only if embeddability in a surface $S$ depends only on $\varepsilon(S)$. For $\varepsilon = 0$ this is not an issue, since the sphere is the only surface with $\varepsilon = 0$. For even $\varepsilon > 0$, however, there are two surfaces of Euler genus $\varepsilon$—one orientable and one non-orientable—and the corresponding classes of graphs embeddable in them do not coincide. (Indeed, large projective-planar grids have unbounded orientable genus [2], while $K_7$ can be embedded in the torus but not in the Klein bottle [6].) Our best hope, therefore, is to characterise embeddability not in a given surface $S$, but in 'some' surface of given Euler genus.

Another flaw in the Naive Conjecture which the reader will have noticed is its reference to cycles: for higher surfaces, even genus-embeddings of 2-connected graphs can have facial walks that are not cycles. (For example, we can embed the graph $G$ of Figure 1 in the torus by running the edge $e = uv$ along a handle added to the sphere to join two triangular faces containing $u$ and $v$, respectively. Then $e$ lies on the boundary of only one face, whose facial walk contains it twice and therefore is not a cycle. Zha [12] constructed for every surface $S$ other than the sphere and the projective plane a 2-connected graph that has a genus-embedding in $S$ but no embedding whose facial walks are all cycles.)

With these two modifications, our conjecture might become the following:

**Revised Conjecture.** *For every integer $\varepsilon \geq 0$, a graph $G$ embeds in a surface of Euler genus at most $\varepsilon$ if and only if it has a family of closed walks that covers every edge at most twice and whose codimension in $\mathcal{C}(G)$ is at most $\varepsilon$.*

However, as noticed already by Mohar [8], this is still not true. In fact, our list of properties of facial cycles—so far: sparseness and large dimension—needs two more additions.

For the first of these, consider the plane graph $A_1$ shown in solid lines in Figure 1. Let $G$ be obtained from $A_1$ by adding the edge $uv$. This graph $G$ is one of the 35 forbidden minors that characterise embeddability in the projective plane (Archdeacon [1]); in particular, $\varepsilon(G) \geq 2$.

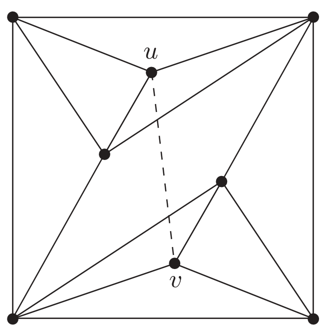


Figure 1: Add the edge $uv$ to obtain a graph $G$ with $\varepsilon(G) \geq 2$

Let $\mathcal{W}$ denote the family of facial walks of $A_1$. The subspace it spans in $\mathcal{C}(G)$ is the cycle space of $A_1$. By (1), and since $G$ has one more edge than $A_1$

but the same number of vertices, we deduce that

$$\dim \mathcal{W} = \dim \mathcal{C}(A_1) = \dim \mathcal{C}(G) - 1.$$

By the Revised Conjecture for $\varepsilon = 1$, this implies that $G$ can be embedded in the projective plane—which it cannot.

To rule out counterexamples such as this, we shall add to our list of requirements on the closed walks in the Revised Conjecture that they cover every edge of $G$—not only at most twice, but also at least once. Since the facial walks in any 2-cell-embedding of a graph have this property, it is certainly an acceptable addition to our list. (In MacLane's theorem no such requirement is needed, because it follows: when $G$ is 2-connected—as we may assume—and the given family of cycles generates all of $\mathcal{C}(G)$, it automatically covers all the edges.)

For our second additional requirement, consider the plane graph $A_2$ shown in Figure 2. Let $G$ be obtained from $A_2$ by identifying the vertices $u$ and $v$. This graph $G$ is another of Archdeacon's 35 forbidden minors for the projective plane, so again $\varepsilon(G) \geq 2$.

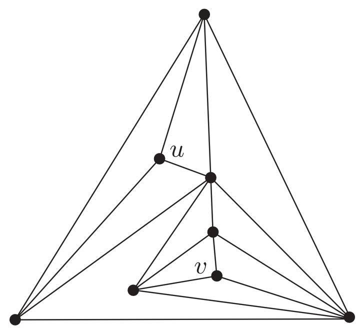


Figure 2: Identify $u$ and $v$ to obtain a graph $G$ with $\varepsilon(G) \geq 2$

As before, the subspace $\mathcal{W}$ spanned in $\mathcal{C}(G)$ by the facial walks of $A_2$ is the cycle space of $A_2$. By (1), and since $G$ has one vertex less than $A_2$ but the same number of edges, we deduce that

$$\dim \mathcal{W} = \dim \mathcal{C}(A_2) = \dim \mathcal{C}(G) - 1.$$

The Revised Conjecture for $\varepsilon = 1$, even if amended by the requirement that $\mathcal{W}$ cover all the edges of $G$ (which it does), thus implies that $G$ can be embedded in the projective plane—which it cannot. But how can we distill from this example a property of the facial cycles of graphs in surfaces that $\mathcal{W}$ fails to have? This question puzzled us for a while, until it occured to us to do the obvious thing: to form the 2-complex corresponding to $\mathcal{W}$ and see why it is not the projective plane. The answer, of course, is that this complex is not a surface at all: it is the pseudo-surface obtained from a sphere by identifying two points.

To rule out this type of counterexample we must therefore require that, for every vertex $v$, no proper subset of those of our given walks that pass through $v$ can combine to a flat neighbourhood of $v$ when we attach 2-cells to these walks. To do this correctly will be a little technical, and we address this task in the next section. Once this is done, we can state and prove our results.

# 4 Statement and discussion of results

Let $W = v_1 e_1 \dots v_n e_n v_1$ be a closed walk in a loopless graph $G$, where the $v_i$ are vertices and the $e_i$ are edges. For a vertex $v$ we call a subsequence $e_{j-1} v_j e_j$ of $W$ with $v_j = v$ (where $e_0 := e_n$) a *pass* of $W$ *through* the vertex $v$. Extending our earlier notation for walks, we write $c(e_{j-1} v_j e_j) := \{e_{j-1}, e_j\}$ if $e_{j-1} \neq e_j$, and $c(e_{j-1} v_j e_j) := \emptyset$ if $e_{j-1} = e_j$.

In order keep track of how often a given walk passes through a given vertex, we shall consider the *family of all passes of $W$ through $v$*, the family $(e_{j-1} v_j e_j)_{j \in J}$ where $J = \{j : v_j = v,\ 1 \leq j \leq n\}$. Similarly, if $\mathcal{W} = (W_i)_{i \in I}$ is a family of walks then the *family of all passes of $\mathcal{W}$ through $v$* is the family $\mathcal{A}(\mathcal{W}, v) := (p_{ij})_{i \in I, j \in J_i}$ where, for each $i$, $(p_{ij})_{j \in J_i}$ is the family of all passes of $W_i$ through $v$.

Recall that a family $\mathcal{F}$ of subsets of $E(G)$ is *sparse* if every edge of $G$ lies in at most two members of $\mathcal{F}$. Similarly, we shall call our family $\mathcal{W}$ of walks *sparse at* an edge $e \in G$ if it covers $e$ at most twice. In view of our discussion in Section 3, we now wish to define a sparseness requirement also at vertices. Let us call $\mathcal{W}$ *sparse at* a vertex $v \in G$ if every non-empty family $\mathcal{F} \subseteq \mathcal{A}(\mathcal{W}, v)$ for which $\sum_{p \in \mathcal{F}} c(p) = \emptyset$ satisfies $\mathcal{F} = \mathcal{A}(\mathcal{W}, v)$.[1] If $\mathcal{W}$ is sparse at all vertices and edges of $G$, we call it a *sparse* family of walks. (For edge sets rather than families of walks we retain our earlier notion of sparseness, meaning sparseness at edges. When this is the intended meaning we shall specifically say that we are speaking about edge sets.)

We can now state our first extension of MacLane's theorem:

**Corollary 2.** *For every integer $\varepsilon \geq 0$, a loopless graph $G$ can be embedded in some surface of Euler genus at most $\varepsilon$ if and only if there is a sparse family of closed walks covering $G$ whose codimension in $\mathcal{C}(G)$ is at most $\varepsilon$.*

For $\varepsilon = 0$, Corollary 2 reduces to MacLane's theorem. This is not immediately obvious: one has to show that a sparse family $\mathcal{B}$ of edge sets of cycles generating $\mathcal{C}(G)$ (as in MacLane's theorem) must be sparse also at vertices when these cycles are interpreted as walks. But this is easy. Indeed, suppose that the edges at $v$ in a non-empty subfamily $\mathcal{F} = (C_1, \dots, C_k)$ of $\mathcal{B}$ sum to zero, but there exists another cycle $C \in \mathcal{B} \setminus \mathcal{F}$ that contains an edge at $v$. Pick edges $uv \in C_1$ and $vw \in C$. If we assume that $G$ is 2-connected (as we may, for a proof of MacLane's theorem), then $G - v$ contains a $u$–$w$ path $P$, and $C' = uPwvu$ is a cycle in $G$. We claim that no set $\mathcal{B}' \subseteq \mathcal{B}$ can sum to $C'$. For since $\mathcal{B}$ is sparse and $\mathcal{F}$ sums to zero at $v$, every edge $e \in \bigcup \mathcal{F}$ at $v$ lies on exactly two $C_i$ but not on any cycle in $\mathcal{B} \setminus \mathcal{F}$; in particular, $e \neq vw$. Hence in order for $\sum \mathcal{B}'$ to contain the edge $uv \in C'$, the family $\mathcal{B}'$ has to meet $\mathcal{F}$ in a proper non-empty subfamily $\mathcal{F}' \subset \mathcal{F}$. But then $\sum \mathcal{F}'$ contains another edge $e \neq uv$ at $v$, which lies neither on $C'$ nor on any cycle in $\mathcal{B}' \setminus \mathcal{F}'$. Hence $e$ lies in $\sum \mathcal{B}'$ but not in $C'$, giving $\sum \mathcal{B}' \neq C'$ as claimed.

Our approach to proving the backward implication of Corollary 2 will be that outlined in Section 3: to each walk in the given sparse family we attach

[1] In plain English: no proper non-empty subfamily of all the passes of $\mathcal{W}$ through $v$ can sum to zero. In our graph $G$ obtained from the graph $A_2$ of Figure 2 by identifying $u$ and $v$, the family of (six) passes of facial walks of $A_2$ through the identified vertex has two such subfamilies: one consisting of the three passes of facial walks of $A_2$ through $u$, the other consisting of the three passes of facial walks of $A_2$ through $v$.

a disc, and prove that the resulting space is a surface. The following theorem spells out this approach explicitly: its backward direction is stronger than that of Corollary 2 in that it allows us to make our *given* walks into face boundaries.

**Theorem 3.** *Let $G = (V, E)$ be a loopless connected graph, $\mathcal{W}$ a family of closed walks in $G$ covering $G$, and $\varepsilon \geq 0$ an integer. Then the following two statements are equivalent:*

(i) *There is a surface $S$ of Euler genus at most $\varepsilon$ in which $G$ can be 2-cell-embedded so that $\mathcal{W}$ is a subfamily of the family of facial walks.*

(ii) *There is a sparse family of closed walks in $G$ that has codimension at most $\varepsilon$ in $\mathcal{C}(G)$ and includes $\mathcal{W}$.*

Given our discussion at the end of Section 3, it may come as a surprise that in the harder implication (ii)→(i) of Theorem 3 we can weaken the premise (ii) considerably while retaining (i) in only slightly weaker form. By our definition of sparseness at vertices, condition (ii) of Theorem 3 requires that no proper subset of all the passes of $\mathcal{W}$ through a vertex $v$ may sum to zero. Since each walk $W \in \mathcal{W}$ can contribute several such passes, the condition becomes much less stringent if we ask the same only for sums of passes that include, for every $W \in \mathcal{W}$, either all the passes of $W$ through $v$ or none. All the same, this weaker condition implies almost the same as statement (i) of Theorem 3: we shall still be able to embed $G$ in a surface of Euler genus at most $\varepsilon$ so that all the subgraphs $G[W]$ with $W \in \mathcal{W}$ become face boundaries. In fact, the embedding will differ from that required by (i) of Theorem 3 only in that the walks $W \in \mathcal{W}$, when we turn them into facial walks, may have the order changed in which their edges are traversed: the set of these edges, as well as their mulitiplicity on $W$ (1 or 2), will remain the same.

How is this possible? Consider a double cover $\mathcal{W}$ of $G$ that is not sparse at vertices. Then a CW-complex constructed by attaching a 2-cell to every walk in $\mathcal{W}$ will not be a surface: it will be a pseudo-surface, with a pinchpoint at every vertex where $\mathcal{W}$ was not sparse. Our claim, therefore, is that by merely traversing the edges of our walks $W$ in a different order we can 'de-pinch' this pseudo-surface, turning it into a proper surface, while keeping its face boundaries both as subgraphs and in terms of edge multiplicities. This is not unremarkable, in view of the fact that such 'de-pinching' was not possible for our second projective-planar example in Section 3.

To make all this precise, let $\mathcal{W}$ be a family of walks in a graph $G = (V, E)$. For a vertex $v$ denote by $\mathcal{W}(v)$ the subfamily of walks containing $v$. We call $\mathcal{W}$ *weakly sparse at* $v$ if every non-empty subfamily $\mathcal{U}$ of $\mathcal{W}(v)$ for which $\sum_{U \in \mathcal{U}} c(U) \cap E(v) = \emptyset$ satisfies $\mathcal{U} = \mathcal{W}(v)$. A family of walks that is sparse at edges as well as weakly sparse at all vertices will be called *weakly sparse*. If $G$ is loopless, then any sparse family of closed walks in $G$ is also weakly sparse, since for every vertex $v$ and every closed walk $W$ we have $c(W) \cap E(v) = \sum_{p \in \mathcal{A}((W),v)} c(p)$.

Finally, call a family $\mathcal{W}' = (\, W' : W \in \mathcal{W}\,)$ of closed walks *similar* to $\mathcal{W}$ if, for every $e \in E(G)$ and every $W \in \mathcal{W}$, the edge $e$ occurs on $W'$ as often as it does on $W$. Thus if $\mathcal{W}'$ is similar to $\mathcal{W}$ then $G[W'] = G[W]$ and $c(W') = c(W)$ for every $W \in \mathcal{W}$, and $\dim \mathcal{W}' = \dim \mathcal{W}$. In particular, if $\mathcal{W}$ is weakly sparse then so is $\mathcal{W}'$—an implication that can fail for 'sparse'.

**Theorem 4.** *Let $G$ be a connected graph, $\mathcal{W}$ a family of closed walks in $G$ covering $G$, and $\varepsilon \geq 0$ an integer. Then the following statements are equivalent:*

(i) *There is a surface of Euler genus at most $\varepsilon$ in which $G$ can be 2-cell-embedded so that the family of facial walks has a subfamily similar to $\mathcal{W}$.*

(ii) *There is a weakly sparse family of closed walks in $G$ that has codimension at most $\varepsilon$ in $\mathcal{C}(G)$ and includes $\mathcal{W}$.*

The graph in Theorem 4 may contain loops: we excluded these in Theorem 3 only because they would have complicated the definition of passes through a vertex, which is implicit in the definition of 'sparse' but not of 'weakly sparse'.

Theorem 4 can be used to strengthen Corollary 2:

**Corollary 5.** *For every integer $\varepsilon \geq 0$, a graph $G$ can be embedded in some surface of Euler genus at most $\varepsilon$ if and only if there is a weakly sparse family of closed walks covering $G$ whose codimension in $\mathcal{C}(G)$ is at most $\varepsilon$.*

For the projective plane it is possible to rephrase our results in terms of cycles rather than arbitrary closed walks, as in MacLane's theorem for the plane. To do so we need the following lemma, which was proved independently by Negami [10] and by Robertson and Vitray [11].

**Lemma 6.** *In any embedding of a 2-connected non-planar graph in the projective plane, all face boundaries are cycles.*

Together with Corollaries 2 and 5 this yields:

**Corollary 7.** *A 2-connected graph is projective-planar if and only if it contains a sparse (equivalently: weakly sparse) family of cycles that covers it and has codimension at most 1 in its cycle space.* □

# 5 The proofs

Let $\mathcal{W}$ be a family of closed walks in a loopless graph $G$ that is sparse at edges. Recall that, for each vertex $v \in G$, we denoted by $\mathcal{A}(\mathcal{W}, v)$ the family of all passes of $\mathcal{W}$ through $v$. As a tool for our proofs, let us define for every vertex $v$ an auxiliary graph $H = H(\mathcal{W}, v)$ with vertex set of $\mathcal{A}(\mathcal{W}, v)$. Its edge set will be a subset of $E(G)$, with incidences defined as follows. Whenever two distinct vertices $p, q$ of $H$ (i.e., passes that are distinct as family members—they may be equal as triples) share an edge $e \in G$, we let $e$ be an edge of $H$ joining $p$ and $q$. If $\mathcal{W}$ contains a pass $p = eve$, we let $e$ be a loop at $p$. Clearly, $H$ has maximum degree at most 2, since a pass $evf$ can be incident only with the edges $e$ and $f$. (For example, if there are three edges $e, f, g$ at $v$ in $G$, and $\mathcal{W}$ contains the passes $evf, fvg, gve$, then these three passes and the three edges $e, f, g$ form a triangle in $H$. As another example, if $\mathcal{W}$ has two passes consisting of the triple $evf$, or one pass $evf$ and another pass $fve$, then these two passes are joined by the pair $\{e, f\}$ of double edges in $H$ and have no other incident edge.)

Note that $\mathcal{W}$ fails to be sparse at the vertex $v$ if and only if $H = H(\mathcal{W}, v)$ contains a cycle $C$ with a vertex set $U \subsetneqq V(H)$, because $\sum_{p \in U} c(p) = 0$ if and only if $H[U]$ is 2-regular. Thus, $\mathcal{W}$ is sparse at $v$ if and only if $H(\mathcal{W}, v)$ is either empty or a single cycle or a forest. If $\mathcal{W}$ is a double cover of $G$, then every $H(\mathcal{W}, v)$ is 2-regular.

We begin with the proof of Theorem 3. For the easier implication, (i)→(ii), we need that the family $\mathcal{W}$ of facial walks of a graph $G$ in a surface $S$ is as close to being linarly independent as any family of edge sets summing to zero can be: that $\dim \mathcal{W} = |\mathcal{W}| - 1$. We shall prove this more generally for arbitrary sparse families of closed walks (indeed of edge sets[2]), which may be of some independent interest.

**Lemma 8.** *Let $G = (V, E)$ be a connected graph, and let $\mathcal{W}$ be a weakly sparse family of non-trivial walks in $G$ covering every edge twice. Then* $\dim \mathcal{W} = |\mathcal{W}| - 1$.

*Proof.* Since $\mathcal{W}$ covers every edge twice we have $\sum_{W \in \mathcal{W}} c(W) = \emptyset$, so $\dim \mathcal{W} < |\mathcal{W}|$. To prove that $\dim \mathcal{W} \geq |\mathcal{W}| - 1$ we show that, in fact, for *every* non-empty proper subfamily $\mathcal{U}$ of $\mathcal{W}$ the family $(\, c(U) : U \in \mathcal{U}\,)$ is linearly independent in the edge space of $G$ over $\mathbb{Z}_2$. Suppose not, i.e., suppose that $\sum_{U \in \mathcal{U}} c(U) = \emptyset$ for some such $\mathcal{U}$.

Pick a walk $W \in \mathcal{W} \setminus \mathcal{U}$. By assumption the walks in $\mathcal{W}$ are non-trivial, so $W$ contains an edge. This edge cannot lie on any walk in $\mathcal{U}$: as $\sum_{U \in \mathcal{U}} c(U) = \emptyset$, this would mean that $\mathcal{U}$ would cover it at least twice, but by assumption all of $\mathcal{W}$ covers it only twice. On the other hand, since $\mathcal{U}$ is non-empty and its walks are non-trivial, $\mathcal{U}$ covers some other edge of $G$. Since $G$ is connected, it therefore has a vertex $v$ that is incident both with an edge covered by $\mathcal{U}$ and with an edge that is covered by $\mathcal{W}$ (by assumption) but not by $\mathcal{U}$. Then the family $\mathcal{U}(v)$ of all walks in $\mathcal{U}$ containing $v$ is a non-empty proper subfamily of the corresponding subfamily $\mathcal{W}(v)$ of $\mathcal{W}$. As

$$\sum_{U \in \mathcal{U}(v)} c(U) \cap E(v) \subseteq \sum_{U \in \mathcal{U}} c(U) = \emptyset, \tag{3}$$

this contradicts our assumption that $\mathcal{W}$ is weakly sparse (at $v$). □

**Proof of Theorem 3.** (i)→(ii) Extend $\mathcal{W}$ to the family $\mathcal{W}'$ of all the facial walks of $G$ in $S$. Since $S$ is locally homeomorphic to the plane, $\mathcal{W}'$ covers every edge of $G$ twice, and elementary topological arguments show that $\mathcal{W}'$ is sparse also at vertices. Hence $\dim \mathcal{W}' = |\mathcal{W}'| - 1$ by Lemma 8. Using (1) and Euler's formula (2), we deduce that

$$\dim \mathcal{C}(G) - \varepsilon = |E(G)| - |V(G)| + 1 - \varepsilon = |\mathcal{W}'| - 1 = \dim \mathcal{W}'$$

as desired.

(ii)→(i). Replacing $\mathcal{W}$ with the extension of $\mathcal{W}$ whose existence is asserted in (ii), we may assume that $\mathcal{W}$ itself is sparse and has codimension at most $\varepsilon$ in $\mathcal{C}(G)$. Let us further assume that the extension was chosen maximal, and begin our proof by showing that $\mathcal{W}$ now is a double cover of $G$.

Suppose not. By assumption, $\mathcal{W}$ covers $G$. The set $F := \sum_{W \in \mathcal{W}} c(W)$ of edges covered only once, therefore, contains all the edges not covered twice; in particular $F \neq \emptyset$. Note also that $F$ lies in $\mathcal{C}(G)$, because every $c(W)$ does; hence every vertex of $(V, F)$ has even degree. Our aim is to find a closed walk

[2]If we had defined sparseness at vertices for families of arbitrary sets of edges rather than just for edge sets of the form $c(U)$, as in the definition of 'weakly sparse', our proof below would prove the lemma for arbitrary sparse families of edge sets.

$W$ in $(V, F)$ such that $\mathcal{W}' := \mathcal{W} \cup \{W\}$ is again sparse; this will contradict our maximal choice of $\mathcal{W}$.

For every vertex $v$ incident with an edge in $F$, consider the auxiliary graph $H(v) := H(\mathcal{W}, v)$ defined at the start of this section. We know that $H(v) \neq \emptyset$, since $\mathcal{W}$ covers $G$. Let us show that $H(v)$ is a forest. Suppose not, and let $U$ be the vertex set of a cycle in $H(v)$. Then $\sum_{u \in U} c(u) = 0$. But this contradicts the sparseness of $\mathcal{W}$ at $v$. Indeed, by assumption $v$ is incident also with an edge $f \in F$, and since $\mathcal{W}$ covers $G$, this edge also lies in some pass of $\mathcal{W}$ through $v$. That pass, however, is not in $U$: as $\mathcal{W}$ covers $f$ only once, it has degree at most 1 in $H(v)$.

The components of $H(v)$, therefore, are paths. For every such path $P$ put $\partial P := \sum_{p \in V(P)} c(p)$; this is a set of two edges in $F \cap E(v)$, and $F \cap E(v)$ is the disjoint union of these 2-sets. Let $C(v)$ be a cycle on $F \cap E(v)$ as its vertex set such that $E(C(v)) \supseteq \{\, \partial P : P \text{ is a component of } H(v) \,\}$. Call the edges in this last set *red*, and the other edges of $C(v)$ *green*; the edges of $C(v)$ thus alternate between red and green. (If $H(v)$ has only one component, then $C(v)$ is a pair of parallel edges, one red and one green.)

To construct our additional walk $W$ in $(V, F)$, we start by picking a vertex $v_1$ of $G$ that is incident with an edge in $F$. Then $H(v_1)$ and $C(v_1)$ are defined. Let $W = v_1 f_1 v_2 f_2 \dots f_{n-1} v_n$ be the unique maximal walk in $(V, F)$ such that every $f_i$ is joined to $f_{i+1}$ in $C(v_{i+1})$ by a green edge and all the $f_i$ are distinct. This walk $W$ can only end in $v_1$ (with the edge that is joined to $f_1$ in $C(v_1)$ by a green edge), so $W$ is closed.

Since $W$ contains each of its edges $f_i$ only once, the extended family $\mathcal{W}' = \mathcal{W} \cup \{W\}$ is again sparse at edges; let us show that it is sparse also at vertices. The passes of $W$ through a vertex $v$ are all triples $evf$ such that $ef$ is a green edge of $C(v)$. Adding these passes as new vertices to $H(v)$, with adjacencies as defined before, turns $H(v)$ into a graph $H'(v)$ that is either a single cycle (if $W$ 'traverses' every green edge of $C(v)$) or a disconnected graph whose components are still paths: $H'(v)$ cannot contain cycles other than a Hamilton cycle, because $C(v)$ is a single cycle. Therefore, as any family $\mathcal{F}$ of passes of $\mathcal{W}'$ through $v$ with $\sum_{p \in \mathcal{F}} c(p) = \emptyset$ induces a cycle in $H'(v)$, this can happen only when $\mathcal{F}$ is the family of *all* passes of $\mathcal{W}'$ through $v$. Thus, $\mathcal{W}'$ is again sparse at vertices, contradicting the maximal choice of $\mathcal{W}$. This completes the proof that $\mathcal{W}$ is a double cover of $G$.

Let us now construct the surface required for statement (i) of the theorem. We shall obtain this as a 2-dimensional CW-complex $C$. To construct $C$, we start with $G$ as its 1-skeleton. As the 2-cells we take disjoint open discs $D_W \subseteq \mathbb{R}^2$, one for each walk $W \in \mathcal{W}$, divide the boundary of $D_W$ into as many segments as $W$ is long, and map consecutive segments homeomorphically to consecutive edges in $W$.

In order for $S := |C|$ to be a surface, we have to check that every point has an open neighbourhood that is homeomorphic to $\mathbb{R}^2$. For points in the interior of 2-cells or edges, this is clear; recall that $\mathcal{W}$ is a double cover. Now consider a vertex $v$ of $G$. Define $H(v)$ as earlier. Since $\mathcal{W}$ is a double cover, $H(v)$ is now 2-regular, and since $\mathcal{W}$ is sparse at $v$ it contains no cycle properly. Hence, $H(v)$ is a cycle. For each pass $p = evf \in V(H(v))$ we let $D(p)$ be a closed disc whose interior lies inside a disc $D_W$ such that $p$ is a pass of $W$, choosing each $D(p)$ so that its boundary contains $v$ and intersects $W$ in one segment contained in $e \cup f$

and meeting both $e$ and $f$. These discs $D(p)$ can clearly be chosen with disjoint interiors for different $p$. Using the elementary fact that the union of two closed discs intersecting in a common segment of their boundaries is again a disc, one easily shows inductively that the interior of the union of all the discs $D(p)$ is an open disc, and hence homeomorphic to $\mathbb{R}^2$. This completes the proof that $S$ is a surface.

Since $C$ is finite, $S$ is compact. Since $G$ is connected, so is $S$. Finally, Euler's formula (2) applied to $C$, together with (1), Lemma 8 (actually, its trivial inequality), and our assumption that $\mathcal{W}$ has codimension at most $\varepsilon$ in $\mathcal{C}(G)$, yields

$$\begin{aligned}\varepsilon(S) &= 2 - (|V(G)| - |E(G) + |\mathcal{W}|)\\ &= (|E(G)| - |V(G)] + 1) - (|\mathcal{W}| - 1)\\ &= \dim \mathcal{C}(G) - \dim \mathcal{W}\\ &\leq \varepsilon\,.\end{aligned}$$

Thus, (i) is proved. □

Theorem 3 implies the forward implication of Corollary 2, as follows.

**Proof of Corollary 2.** The backward direction follows immediately from that of Corollary 5.

For the forward direction, let $G$ and $\varepsilon$ be such that $G$ embeds in a surface of Euler genus at most $\varepsilon$. Our aim is to find a certain family of closed walks of codimension at most $\varepsilon$, so there is no loss of generality in choosing $\varepsilon$ minimal, i.e., in assuming that $\varepsilon = \varepsilon(G)$. Let $G_1, \dots, G_n$ be the components of $G$. For each $i = 1, \dots, n$ choose a genus-embedding $G_i \hookrightarrow S_i$. These are 2-cell-embeddings, and by genus additivity we have $\varepsilon_1 + \dots + \varepsilon_n = \varepsilon$ for $\varepsilon_i := \varepsilon(S_i) = \varepsilon(G_i)$. For each $i$ let $\mathcal{W}_i$ be the family of facial walks of $G_i$ in $S_i$. By Theorem 3, the $\mathcal{W}_i$ are sparse and have codimension at most $\varepsilon_i$ in $\mathcal{C}(G_i)$: as $\mathcal{W}_i$ already covers every edge of $G_i$ twice, it cannot be extended to a larger sparse family. Since the $G_i$ are vertex-disjoint, $\mathcal{W} := \mathcal{W}_1 \cup \dots \cup \mathcal{W}_n$ is again sparse, and it has codimension at most $\varepsilon_1 + \dots + \varepsilon_n = \varepsilon$ in $\mathcal{C}(G)$, since $\mathcal{C}(G)$ is the direct sum of the spaces $\mathcal{C}(G_i)$. □

Next, we prove Theorem 4. The core of the proof (from Theorem 3) is the following 'de-pinching' lemma. It employs a standard trick from surface surgery to dissolve pinchpoints, which was also used by Edmonds [5]. In our terminology, it turns a weakly sparse family of walks into a sparse one:

**Lemma 9.** *For every weakly sparse family $\mathcal{W}$ of closed walks in a connected loopless graph $G$ there exists a sparse family $\mathcal{W}'$ similar to $\mathcal{W}$.*

*Proof.* Among all the families $\mathcal{W}' = \{\, W' : W \in \mathcal{W} \,\}$ similar to $\mathcal{W}$ choose one with $\sum_{v \in V(G)} \gamma(v)$ minimum, where $\gamma(v)$ denotes the number of components of $H(\mathcal{W}', v)$. We show that $\mathcal{W}'$ is sparse.

Suppose not. Being similar to $\mathcal{W}$, the family $\mathcal{W}'$ is again weakly sparse, and in particular sparse at edges. So it must fail to be sparse at some vertex $v$. Then $H' := H(\mathcal{W}', v)$ contains a cycle $C$ that is a component of $H'$ but not equal to $H'$. Since $\mathcal{W}'$ is weakly sparse at $v$, one of the vertices of $C$ must be a pass $p = evf$ of a walk $W' \in \mathcal{W}'$ which also contains a pass $p' = e'vf'$ that is a vertex in another component $C' \neq C$ of $H'$. Choose these passes so that $W'$ has a subwalk $vf \ldots e'v$ not containing $e$ or $f'$. Let $W''$ be the closed walk obtained from $W'$ by reversing this subwalk (Figure 3), and let $\mathcal{W}''$ be obtained from $\mathcal{W}'$ by replacing $W'$ with $W''$. Clearly, $W''$ is again a closed walk, and $\mathcal{W}''$ is similar to $\mathcal{W}'$ and hence to $\mathcal{W}$.

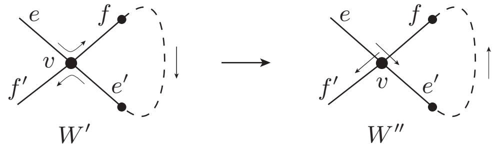


Figure 3: Turning $W'$ into $W''$ by reversing the segment $vf \ldots e'v$

For vertices $u \neq v$ of $G$ we have $H(\mathcal{W}'', u) = H(\mathcal{W}', u)$, so $\gamma(u)$ remains unchanged. At $v$, however, $\gamma(v)$ decreases, contradicting the choice of $\mathcal{W}'$. Indeed, $H'' := H(\mathcal{W}'', v)$ arises from $H'$ by the replacement of $p = evf \in V(C)$ and $p' = e'vf' \in V(C')$ with two new vertices, $q := eve'$ and $q' := fvf'$, and redefining the incidences for the edges $e, f, e', f' \in E(H') = E(H'')$ accordingly. As one easily checks (see Figure 4), this has the effect of merging the components $C$ and $C'$ of $H'$ into one new component, leaving the other components of $H'$ intact. Thus, the components of $H''$ are those of $H'$ other than $C$ and $C'$, plus one new component arising from $(C - p) \cup (C' - p')$ by adding the new vertex $q$ incident with $e$ and $e'$ and the new vertex $q'$ incident with $f$ and $f'$ (leaving the other incidences of $e, e', f, f'$ in $H''$ as they were in $H'$). $\square$

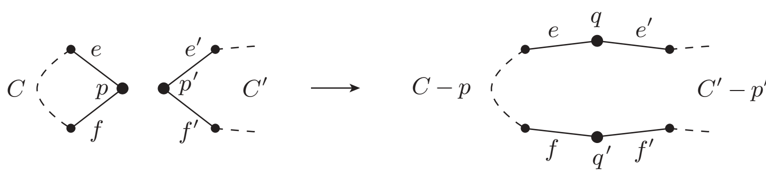


Figure 4: Merging the components $C$ and $C'$ of $H'$ to form $H''$

**Proof of Theorem 4.** Subdividing every loop once, we may assume that $G$ has no loops.

To prove the implication (i)$\to$(ii), consider an embedding of $G$ as in (i). Let $\mathcal{W}'$ denote the subfamily of the family of facial walks that is similar to $\mathcal{W}$. By Theorem 3, $\mathcal{W}'$ can be extended to a sparse family of closed walks in $G$ that has codimension at most $\varepsilon$ in $\mathcal{C}(G)$. Replacing $\mathcal{W}'$ by $\mathcal{W}$ in this family, we obtain an extension of $\mathcal{W}$ that is still weakly sparse, and hence satisfies (ii).

For a proof of the implication (ii)$\to$(i), let us now assume (ii). Consider the weakly sparse family $\widetilde{\mathcal{W}} \supseteq \mathcal{W}$ provided by (ii), and use Lemma 9 to turn it into a sparse family $\widetilde{\mathcal{W}}'$ similar to $\widetilde{\mathcal{W}}$. Then $\widetilde{\mathcal{W}}'$ has a subfamily $\mathcal{W}'$ similar to $\mathcal{W}$. By similarity, $\widetilde{\mathcal{W}}'$ has the same codimension in $\mathcal{C}(G)$ as $\widetilde{\mathcal{W}}$, at most $\varepsilon$. By Theorem 3, there is a surface $S$ of Euler genus at most $\varepsilon$ in which $G$ can be 2-cell-embedded so that the walks in $\widetilde{\mathcal{W}}'$ are among the facial walks. This embedding satisfies (i), because $\widetilde{\mathcal{W}}'$ contains $\mathcal{W}'$. $\square$

Corollary 5 now follows from Theorem 4 and the forward direction of Corollary 2 (which we deduced from Theorem 3):

**Proof of Corollary 5.** For the forward direction, let $G$ be a graph embeddable in a surface of Euler genus at most $\varepsilon$. Subdivide every loop once to obtain a loopless graph $\dot{G}$. Clearly, the embedding of $G$ defines one of $\dot{G}$. By the forward direction of Corollary 2, there is a sparse family $\mathcal{W}$ of closed walks in $\dot{G}$ that covers $\dot{G}$ and whose codimension in $\mathcal{C}(\dot{G}) \doteq \mathcal{C}(G)$ is at most $\varepsilon$. Our task is to show that every non-trivial walk $W \in \mathcal{W}$ defines a walk in $G$, i.e., contains no pass of the form $eve$ through a subdividing vertex $v$.

Suppose $W$ does contain such a pass $eve$, and let $f$ be the other edge of $\dot{G}$ at $v$. Then the family $\mathcal{F} = \{eve\}$ satisfies $\sum_{p \in \mathcal{F}} c(p) = 0$, and since $\mathcal{W}$ covers $f$, $\mathcal{F}$ is a proper subfamily of $\mathcal{A}(\mathcal{W}, v)$. This contradicts the fact that $\mathcal{W}$ is sparse at $v$.

For a proof of the backward direction, let $\mathcal{W}$ be a weakly sparse family of closed walks in $G$ that covers $G$ and has codimension at most $\varepsilon$ in $\mathcal{C}(G)$. If $G$ has components $G_1, \ldots, G_k$, say, write $\mathcal{W}_i$ for the subfamily of walks contained in $G_i$, and $\varepsilon_i$ for the codimension of $\mathcal{W}_i$ in $\mathcal{C}(G_i)$. Then $\varepsilon(G_i) \leq \varepsilon_i$, by (ii)$\to$(i) of Theorem 4. Moreover, $\sum_{i=1}^{k} \varepsilon_i \leq \varepsilon$, since $\mathcal{C}(G)$ is the direct sum of the spaces $\mathcal{C}(G_i)$. Hence, by genus additivity,

$$\varepsilon(G) = \sum_{i=1}^{k} \varepsilon(G_i) \leq \sum_{i=1}^{k} \varepsilon_i \leq \varepsilon\,.$$

Thus, $G$ can be embedded in a surface of Euler genus at most $\varepsilon$. $\square$


## Acknowledgement

We thank Bojan Mohar for updating us at the start of this project on what was known on the subject.


## References


[1] D. Archdeacon, *A Kuratowski theorem for the projective plane*, J. Graph Theory **5** (1981), 243–246.

[2] L. Auslander, T.A. Brown, and J.W.T. Youngs, *The imbedding of graphs in manifolds*, J. Math. and Mech. **12** (1963), 629–634.

[3] H. Bruhn and R. Diestel, *Embeddability and duality in higher surfaces*, in preparation.

[4] R. Diestel, *Graph theory* (3rd edition), Springer-Verlag, 2005.

[5] J. Edmonds, *On the surface duality of linear graphs*, J. Res. Natl. Bur. Stand., Sect. B **69** (1965), 121–123.

[6] P. Franklin, *A six color problem*, J. Math. Massachusetts **13** (1934), 363–369.

[7] S. MacLane, *A combinatorial condition for planar graphs*, Fund. Math. **28** (1937), 22–32.

[8] B. Mohar, *Generalization of MacLane's characterization of planar graphs*, Congr. Numerantium **56** (1987), 231–236.

[9] B. Mohar and C. Thomassen, *Graphs on surfaces*, Johns Hopkins, 2001.

[10] S. Negami, *Re-embedding of projective-planar graphs*, J. Combin. Theory (Series B) **44** (1988), 276–299.

[11] N. Robertson and R. Vitray, *Representativity of surface embeddings*, Paths, Flows, and VLSI-Layout (H.J. Prömel B. Korte, L. Lovász and A. Schrijver, eds.), Algorithms and Combinatorics, vol. 9, Springer-Verlag, 1990, pp. 293–328.

[12] X. Zha, *On minimum-genus embeddings*, Disc. Math. **149** (1996), 261–278.

Henning Bruhn
`<hbruhn@gmx.net>`
Laboratoire LEIBNIZ
INPG
46 avenue Felix Viallet
38031 Grenoble
France

Reinhard Diestel
`<diestel@math.uni-hamburg.de>`
Mathematisches Seminar
Universität Hamburg
Bundesstraße 55
20146 Hamburg
Germany

Version 12.6.2006