Accepted for publication in
Communications in Mathematical Physics

ROMAN KOTECKÝ1 ALAN D. SOKAL2 JAN SWART3

ENTROPY-DRIVEN PHASE TRANSITION IN LOW-TEMPERATURE ANTIFERROMAGNETIC POTT S MODELS

Abstract. We prove the existence of long-range order at sufficiently low temperatures, including zero temperature, for the three-state Potts antiferromagnet on a class of quasi-transitive plane quadrangulations, including the diced lattice. More precisely, we show the existence of (at least) three infinite-volume Gibbs measures, which exhibit spontaneous magnetization in the sense that vertices in one sublattice have a higher probability to be in one state than in either of the other two states. For the special case of the diced lattice, we give a good rigorous lower bound on this probability, based on computer-assisted calculations that are not available for the other lattices.

Keywords: Antiferromagnetic Potts model, proper coloring, plane quadrangulation, phase transition, diced lattice.

MSC 2000 Subject Classification: Primary 82B20; Secondary 05C15, 05C63, 60K35.

Contents

1. Introduction and main results 2
1.1. Introduction 2
1.2. Statement of the results 5
1.3. Discussion 9
1.4. Some further open problems 13
1.5. Plan of this paper 14
2. Main structure of the proofs 14
2.1. Contour model 14
2.2. The basic lemmas 15
2.3. Proof of the main theorems, given the basic lemmas 18
3. The zero-temperature case 20
3.1. Contour model for zero temperature 20
3.2. Bounds on contours for zero temperature 22
3.3. Long-range dependence for zero temperature 25

1 Charles University, Prague, Czech Republic, and University of Warwick, United Kingdom, R.Kotecky@warwick.ac.uk
2 Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA, and Department of Mathematics, University College London, Gower Street, London WC1E 6BT, United Kingdom, sokal@nyu.edu
3 Institute of Information Theory and Automation of the ASCR (ÚTIA), Pod vodárenskou věží 4, 18208 Praha 8, Czech Republic, swart@utia.cas.cz
1. Introduction and main results

1.1. Introduction. We are interested here in the three-state antiferromagnetic Potts model on a class of infinite plane quadrangulations. Recall that a graph embedded in the plane is called a quadrangulation if all its faces are quadrilaterals (i.e., have four vertices and four edges). Some examples of infinite plane quadrangulations are drawn in Figure 1; these include the square lattice $\mathbb{Z}^2$ (with nearest-neighbor edges) and the so-called diced lattice.

On the square lattice, the three-state Potts antiferromagnet at zero temperature can be mapped into a special case of the six-vertex model that admits an exact (but nonrigorous) solution [5, section 8.13]. This model is therefore believed to be critical at zero temperature but disordered for any positive temperature.

On the diced lattice, by contrast, a proof was outlined in [47] showing that the three-state Potts antiferromagnet has a phase transition at nonzero temperature and has long-range order at all sufficiently low temperatures (including zero temperature). In the present paper, we present the details of this proof and we extend the result to a large class of quasi-transitive quadrangulations, including some hyperbolic lattices.

To explain the class of lattices that we can cover, let us start by observing that a quadrangulation is a connected bipartite graph $G = (V, E)$, so that the vertex set has a canonical bipartition $V = V_0 \cup V_1$. We may view the two sublattices $V_0$ and $V_1$ as graphs in their own right by connecting vertices along the diagonals of

1 In this paper we restrict attention to nondegenerate quadrangulations, i.e. each face has four distinct vertices and four distinct edges. Some discussion of degenerate plane quadrangulations (in the case of finite graphs) can be found in [39].

2 See e.g. the discussion below formula (2.8) in [58]. See also [24, 65] for Monte Carlo data supporting these beliefs.
Figure 1. Four quasi-transitive quadrangulations with their sub-lattices $G_0$ (open circles joined by dashed red edges) and $G_1$ (filled circles joined by solid green edges). In these examples, the sub-lattice $G_0$ is (a) the square lattice, (b) the triangular lattice, (c) the union-jack lattice, and (d) the hyperbolic lattice with Schl"afli symbol $\{3, 7\}$. Note that in (b)–(d), $G_0$ is a triangulation. In (a) and (b), the quadrangulations $G$ are, respectively, the square lattice and the diced lattice.

the quadrilateral faces of the original lattice: this yields graphs $G_0 = (V_0, E_0)$ and $G_1 = (V_1, E_1)$ as shown in Figure 1. Note that $G_0$ and $G_1$ are duals of each other,
Figure 2. The two sublattices $G_0$ and $G_1$ for each of the four lattices from Figure 1. The graph $G_0$ (red open circles and dashed edges) is dual to the graph $G_1$ (green filled circles and solid edges).

i.e. each face of $G_0$ contains a unique vertex of $G_1$, and vice versa; and each edge of $G_0$ crosses a unique edge of $G_1$, and vice versa (see Figure 2). Conversely, given any dual pair of (finite or infinite) graphs $G_0 = (V_0, E_0)$ and $G_1 = (V_1, E_1)$ embedded in the plane, we can form a plane quadrangulation $G = (V, E)$ by setting $V = V_0 \cup V_1$ and placing an edge between each pair of vertices $v \in V_0$ and $w \in V_1$ where $w$ lies in a face of $G_0$ that has $v$ on its boundary (or equivalently vice versa). The main assumption that we will make in this paper is that one sublattice (say, $G_0$)
is a triangulation. In particular, our proofs cover the lattices shown in (b)–(d) of Figure 1, but not the square lattice (a).

To explain the nature of the phase transition, note that ground states of the three-state Potts antiferromagnet are simply proper three-colorings of the lattice. On any bipartite lattice, we may construct special ground states by coloring one sublattice (say, \( V_0 \)) in one color and using the other two colors to color the other sublattice in any possible way. Note that in this way, the second sublattice carries all the entropy. Of course, the special ground states in which the first sublattice uses only one color are atypical of the Gibbs measure, even at zero temperature. Nevertheless, the underlying idea applies more generally: there may be a preference for the first sublattice to be colored mostly in one color because this increases the freedom of choice of colors on the other sublattice. Otherwise put, integrating out the colors on the second sublattice may induce an effective ferromagnetic interaction on the first sublattice. If this effective interaction is strong enough, it may result in long-range order on the first sublattice. We call this an entropy-driven phase transition. In [47] a proof was sketched along these lines that an entropy-driven transition indeed occurs on the diced lattice. Here we will present the details of this proof and extend it to a large class of plane quadrangulations in which one sublattice is a triangulation. The extension uses a variant of the Peierls argument that works whenever the Peierls sum is finite (even if it is not small), followed by a random-cluster argument.

In all the cases handled in this paper, there is a strong asymmetry between the two sublattices, so that it is entropically favorable to ferromagnetically order the triangulation (\( G_0 \)) and place the entropy on the other sublattice (\( G_1 \)). By contrast, in the square lattice, where no finite-temperature phase transition is believed to occur, the two sublattices are isomorphic. It is therefore natural to ask whether asymmetry is a necessary and/or sufficient condition for the existence of a finite-temperature phase transition. This is a subtle question, and we discuss it further in Section 1.3 below.

1.2. Statement of the results. Let us now formulate our results precisely. We first need to define more precisely the class of graphs we will be considering. We quickly review here the essential definitions; a more thorough summary of the needed theory of infinite graphs can be found in the Appendix.

A graph \( G = (V, E) \) is called locally finite if every vertex has finitely many neighbors; in this paper we will consider only locally finite graphs. If \( G \) has at least \( k + 1 \) vertices, then \( G \) is called \( k \)-connected if one needs to remove at least \( k \) vertices to disconnect it. A graph \( G \) is said to have one end if after the removal of finitely many edges, there is exactly one infinite connected component; note that this implies in particular that \( G \) is infinite.

---

3 The method for encoding such entropy costs in terms of certain Peierls contours was suggested already in [46], but in that paper it led to a proof of the transition only for some toy models including the three-state Potts antiferromagnet on the “decorated cubic lattice”.
A graph is said to be planar if it can be drawn in the plane \( \mathbb{R}^2 \) with vertices represented by distinct points and edges represented by closed continuous arcs joining their endvertices, mutually disjoint except possibly at their endpoints. A plane graph is a planar graph with a given embedding in the plane. An embedding of a connected graph in the plane is called edge-accumulation-point-free (or EAP-free for short) if there are no points in the plane with the property that each neighborhood of the point intersects infinitely many edges. An EAP-free embedding divides the plane into connected open sets called faces. The boundary of a face is either a finite cycle or a two-way-infinite path.\(^4\)

Consider an EAP-free embedding of a 3-connected graph \( G \). Then one may define a dual graph \( G^* = (V^*, E^*) \) whose vertex set \( V^* \) is the set of faces of \( G \) and with edge set \( E^* \equiv E \), where by definition two faces of \( G \) (or equivalently two vertices of \( G^* \)) are linked by an edge \( e^* \in E^* \) if and only if the corresponding edge \( e \in E \) lies in the border of both faces. Clearly \( G^* \) is locally finite if and only if every face of \( G \) is bounded by a finite cycle. In this case we can embed \( G^* \) in an EAP-free way in the plane such that each face of \( G \) contains exactly one vertex of \( G^* \) and each edge of \( G \) is crossed by exactly one edge of \( G^* \), and vice versa; in particular, \( G \) is also the dual of \( G^* \). We say that \( G \) is a triangulation (resp. quadrangulation) if every face of \( G \) is bounded by a triangle (resp. quadrilateral), or equivalently if each vertex of \( G^* \) has degree 3 (resp. 4). It turns out that triangulations and quadrangulations, defined in this way, are always graphs with at most one end.\(^5\)

The final set of definitions we need concerns some form of “translation invariance” of our lattices. An automorphism of a graph \( G = (V, E) \) is a bijection \( g: V \rightarrow V \) that preserves the graph structure. Two vertices \( u, v \in V \) are of the same type if there exists an automorphism that maps \( u \) into \( v \). This relation partitions the vertex set \( V \) into equivalence classes called types. The graph \( G \) is called vertex-transitive if there is just one equivalence class, and vertex-quasi-transitive if there are finitely many equivalence classes. Edge-transitivity and (for plane graphs) face-transitivity are defined similarly. The corresponding forms of quasi-transitivity are

\(^4\) The “one-way-infinite ladder” plane graph with vertices at points \( (0, 1/n) \) and \( (1, 1/n) \) \( \{n = 1, 2, \ldots \} \) and straight-line edges joining the pairs \( \{(0, 1/n), (1, 1/n)\}, \{(0, 1/n), (0, 1/(n+1))\} \) and \( \{(1, 1/n), (1, 1/(n+1))\} \) is an example of a locally finite plane graph that is not EAP-free. Indeed, each point \( (t, 0) \) with \( t \in [0, 1] \) is an edge accumulation point.

\(^5\) By “two-way-infinite path” (or “double ray”) we mean simply a graph that is isomorphic to \( \mathbb{Z} \) with nearest-neighbor edges. In an EAP-free embedding, a two-way-infinite path has no accumulation points in the finite plane \( \mathbb{R}^2 \), but both of its outgoing rays tend to infinity. Therefore, in the sphere \( S = \mathbb{R}^2 \cup \{\infty\} \), the closure of a two-way infinite path is homeomorphic to a circle \( S^1 \).

\(^6\) A more general definition, which allows for multiple ends, is discussed in Section A.2 of the Appendix. Briefly, we say that a locally finite 3-connected graph \( G \) is a triangulation (resp. quadrangulation) if \( G \) has an abstract dual in which each vertex has degree 3 (resp. 4); see Section A.2 for the definition of abstract duals. With this more general definition, it can be shown that a triangulation (resp. quadrangulation) has an EAP-free embedding in the plane if and only if it is finite or has one end: see Proposition A.13 in Section A.3.
all equivalent (see Lemma A.1 in the Appendix), which is why we simply talk about quasi-transitivity without specifying whether in the vertex-, edge- or face-sense.

We will study the three-state antiferromagnetic Potts model on plane quadrangulations $G$, constructed from mutually dual sublattices $G_0$ and $G_1$, such that $G_0$ is a locally finite 3-connected quasi-transitive triangulation with one end.\(^\text{7}\) Note that quasi-transitivity refers only to the structure of $G_0$ (or $G_1$ or $G$) as an abstract graph (i.e., without reference to any embedding). It turns out [3, Theorem 4.2] (see also [67, Theorem 1]) that any locally finite 3-connected quasi-transitive planar graph with one end can be periodically embedded in either the Euclidean or hyperbolic plane, i.e., so that the automorphisms of $G$ correspond to a discrete subgroup of the group of isometries of the embedding space. But we do not use this fact anywhere in this paper.

Let us now define the $q$-state Potts antiferromagnet on an arbitrary infinite graph $G = (V, E)$, for an arbitrary positive integer $q$. The state space is the set

$$S := \{\sigma = (\sigma_v)_{v \in V} : \sigma_v \in [q] \ \forall v \in V\},$$

(1.1)

where we have used the shorthand notation $[q] = \{1, 2, \ldots, q\}$. We also let

$$S_g := \{\sigma \in S : \sigma_u \neq \sigma_v \ \forall \{u, v\} \in E\}$$

(1.2)

denote the set of proper $q$-colorings of $G$. We sometimes use the terms “spin configuration” for $\sigma \in S$ and “ground-state configuration” for $\sigma \in S_g$. For each finite subset $\Lambda \subset V$ we let

$$\partial \Lambda := \{v \in V \setminus \Lambda : \{v, u\} \in E \text{ for some } u \in \Lambda\}$$

(1.3)

denote the external boundary of $\Lambda$. For any boundary condition $\tau : V \to [q]$ and any spin configuration $\sigma : \Lambda \to [q]$ on $\Lambda$, we define the Hamiltonian of $\sigma$ under the boundary condition $\tau$ by

$$H_{\Lambda}(\sigma | \tau) := \sum_{u, v \in \Lambda} \delta_{\sigma_u, \sigma_v} + \sum_{\{u, v\} \in E} \delta_{\sigma_u, \tau_v}$$

(1.4)

where $\delta_{\sigma_u, \sigma_v}$ is the Kronecker delta, i.e.

$$\delta_{ab} = \delta(a, b) = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b \end{cases}$$

(1.5)

For $\beta \in [0, \infty)$, we define the Gibbs measure in volume $\Lambda$ with boundary condition $\tau$ at inverse temperature $\beta$:

$$\mu_{\Lambda, \beta}^{\tau}(\sigma) := \frac{1}{Z_{\Lambda, \beta}^{\tau}} \exp[-\beta H_{\Lambda}(\sigma | \tau)].$$

(1.6)

\(^7\) The quasi-transitivity of $G_0$ implies that also $G_1$ is quasi-transitive; see Theorem A.3(v) in the Appendix. It is not hard to see that now also $G$ must be quasi-transitive.
For $\beta = \infty$, we define
\begin{equation}
\mu_{\Lambda,\infty}(\sigma) := \lim_{\beta \to \infty} \mu_{\Lambda,\beta}(\sigma).
\end{equation}
That is, $\mu_{\Lambda,\infty}$ is the uniform distribution on configurations $\sigma$ that minimize $H_\Lambda(\sigma | \tau)$. [Note that for some $\tau$ this minimum energy might be strictly positive, i.e. there might not exist proper colorings of $\Lambda \cup \partial \Lambda$ that agree with $\tau$ on $\partial \Lambda$.] Of course, these definitions actually depend on $\tau$ only via the restriction $\tau_{\partial \Lambda} := (\tau_u)_{u \in \partial \Lambda}$ of $\tau$ to $\partial \Lambda$.

We then define infinite-volume Gibbs measures in the usual way through the Dobrushin–Lanford–Ruelle (DLR) conditions [30], i.e., we say that a probability measure $\mu$ on $S$ is an infinite-volume Gibbs measure for the $q$-state antiferromagnetic Potts model at inverse temperature $\beta \in [0, \infty]$ if for each finite $\Lambda \subset V$ its conditional probabilities satisfy
\begin{equation}
\mu(\sigma_{\Lambda} | \sigma_{V \setminus \Lambda} = \tau_{V \setminus \Lambda}) = \mu_{\Lambda,\beta}(\sigma_{\Lambda}) \quad \text{for } \mu\text{-a.e. } \tau.
\end{equation}

For the remainder of this paper we specialize to $q = 3$. Here is our main result:

**Theorem 1.1. (Gibbs state multiplicity and positive magnetization)** Let $G = (V,E)$ be a quadrangulation of the plane, and let $G_0 = (V_0,E_0)$ and $G_1 = (V_1,E_1)$ be its sublattices, with edges drawn along the diagonals of quadrilaterals. Assume that $G_0$ is a locally finite 3-connected quasi-transitive triangulation with one end. Then there exist $\beta_0, C < \infty$ and $\epsilon > 0$ such that for each inverse temperature $\beta \in [\beta_0, \infty]$ and each $k \in \{1,2,3\}$, there exists an infinite-volume Gibbs measure $\mu_{k,\beta}$ for the 3-state Potts antiferromagnet on $G$ satisfying:

(a) For all $v_0 \in V_0$, we have $\mu_{k,\beta}(\sigma_{v_0} = k) \geq \frac{1}{3} + \epsilon$.
(b) For all $v_1 \in V_1$, we have $\mu_{k,\beta}(\sigma_{v_1} = k) \leq \frac{1}{3} - \epsilon$.
(c) For all $\{u,v\} \in E$, we have $\mu_{k,\beta}(\sigma_u = \sigma_v) \leq Ce^{-\beta}$.

In particular, for each inverse temperature $\beta \in [\beta_0, \infty]$, the 3-state Potts antiferromagnet on $G$ has at least three distinct extremal infinite-volume Gibbs measures.

**Remarks.** 1. The bound (c) shows in particular that the zero-temperature Gibbs measure $\mu_{k,\infty}$ is supported on ground states.

2. Any subsequential limit as $\beta \to \infty$ of the measures $\mu_{k,\beta}$ with $\beta < \infty$ also satisfies the bounds (a)–(c). Therefore, there exist zero-temperature Gibbs measures with these properties that are limits of finite-temperature Gibbs measures with these properties.

To see that this is a nontrivial property, consider on the square lattice $\mathbb{Z}^2$ the configuration $\tau \in S$ of the 3-state Potts antiferromagnet defined by $\tau_{(i,j)} = 1 + (i + j \bmod 3)$. Then $\tau \in S_g$ is a ground-state configuration such that its restriction to any row and any column, suitably shifted, is the sequence $(\ldots, 1,2,3,1,2,3,\ldots)$. Since for any finite $\Lambda \subset \mathbb{Z}^2$ there is precisely one ground-state configuration that agrees with $\tau$ on $\mathbb{Z}^2 \setminus \Lambda$, namely $\tau$ itself, we see that the Dirac measure $\delta_\tau$ is a zero-temperature infinite-volume Gibbs measure. But this measure is not a limit of
positive-temperature Gibbs measures: the reason can be traced to the fact that at 
\( \beta < \infty \) with boundary condition \( \tau \) in a sufficiently large volume \( \Lambda \), there exists a 
configuration \( \bar{\tau} \) [namely, \( \bar{\tau}(i,j) = 1 + (i + j \mod 2) \)] such that we can replace \( \tau \) on the 
internal boundary of \( \Lambda \) by \( \bar{\tau} \) at an energetic cost of order \( |\partial \Lambda| \), while gaining a bulk 
entropic advantage (with the boundary condition \( \tau \), one can color \( \Lambda \) with 1 on one 
sublattice and arbitrarily 2 or 3 on the other sublattice).

3. We construct the infinite-volume Gibbs measures \( \mu_{\beta,k} \) as subsequential limits 
of finite-volume Gibbs measures. We expect that there is no need to go to a subse-
quence and that our approximation procedure yields extremal infinite-volume Gibbs 
measures, but we have not proven either of these assertions.

In the special case where \( G \) is the diced lattice, we have a good explicit bound on 
the probabilities in Theorem 1.1(a,b):

**Theorem 1.2. (Quantitative bound for the diced lattice)** Let \( G = (V,E) \) 
be the diced lattice and let \( G_0 = (V_0,E_0) \) and \( G_1 = (V_1,E_1) \) be its triangular and 
and hexagonal sublattices, respectively. Then there exists \( C < \infty \) such that for each 
inverse temperature \( \beta \in [0,\infty] \) and each \( k \in \{1,2,3\} \), there exists an infinite-volume 
Gibbs measure \( \mu_{k,\beta} \) for the 3-state Potts antiferromagnet on \( G \) satisfying:

(a) For all \( v_0 \in V_0 \), we have \( \mu_{k,\beta}(\sigma_{v_0} = k) \geq 0.90301 - Ce^{-\beta} \).

(b) For all \( v_1 \in V_1 \), we have \( \mu_{k,\beta}(\sigma_{v_1} = k) \leq 0.14549 + Ce^{-\beta} \).

(c) For all \( \{u,v\} \in E \), we have \( \mu_{k,\beta}(\sigma_u = \sigma_v) \leq Ce^{-\beta} \).

The lower bound 0.90301 should be compared with the estimated zero-temperature 
value 0.957597 ± 0.000004 from Monte Carlo simulations \[47\].

1.3. **Discussion.** The phase diagram of non-attractive (i.e., non-ferromagnetic) 
spin systems is generally harder to predict than for attractive (ferromagnetic) spin 
systems, and may sometimes depend subtly on the microscopic details of the model. 
In particular, this is true for the two-dimensional 3-state Potts antiferromagnet, for 
which we have shown that it has a phase transition at positive temperature on the 
diced lattice, while no such phase transition is believed to occur on the square lattice — 
even though both lattices are bipartite and are in fact plane quadrangulations.

The existence of a positive-temperature transition in the diced-lattice model was 
a surprise when it was first discovered \[47\], for the following reason: Some two-
dimensional antiferromagnetic models at zero temperature have the property that 
they can be mapped exactly onto a “height” model (in general vector-valued) \[65,38\]. 
In such cases one can argue heuristically that the height model must always be in 
either a “smooth” (ordered) or a “rough” (massless) phase; correspondingly, the 
underlying zero-temperature spin model should either be ordered or critical, never 
disordered. Experience teaches us (or at least seemed to teach us) that the most

---

8 The value \( M_0 = 0.936395 \pm 0.000006 \) reported in \[47\] is the spontaneous magnetization in the 
hypertetrahedral representation, i.e. \( M_0 = \mu_{1,\infty}(\sigma_v = 1) - \frac{1}{2}\mu_{1,\infty}(\sigma_v \neq 1) \).
common case is criticality\(^9\). In particular, when the \(q\)-state zero-temperature Potts antiferromagnet on a two-dimensional periodic lattice admits a height representation, one ordinarily expects that model to have a zero-temperature critical point. This prediction is confirmed (at least non-rigorously) in most heretofore-studied cases: 2-state (Ising) triangular \([7, 57]\), 3-state square-lattice \([58, 43, 11, 65]\), 3-state kagome \([36, 45]\), 4-state triangular \([55]\), and 4-state on the line graph of the square lattice \([44, 45]\). Indeed, before the work of \([47]\), no exceptions were known.

It was furthermore observed in \([47]\) that the height mapping employed for the 3-state Potts antiferromagnet on the square lattice \([65]\) carries over unchanged to any plane quadrangulation. One would therefore have expected the 3-state Potts antiferromagnet to have a zero-temperature critical point on every periodic plane quadrangulation. The example of the diced lattice showed that this is not the case; and the results of the present paper provide further counterexamples. Clearly, the mere existence of a height representation does not guarantee that the model will be critical. Indeed, criticality may well be an exception — corresponding to cases with an unusual degree of symmetry — rather than the generic case.

The mechanism behind all these transitions is what we have called an “entropy-driven phase transition”: namely, ordering on one sublattice increases the entropy available to the other sublattice; or said in a different way, integrating out the spins on the second sublattice induces an effective ferromagnetic interaction on the first sublattice. If this effective interaction is strong enough, it may result in long-range order. Such a phase transition can therefore occur in principle in any antiferromagnetic model on any bipartite lattice\(^10\) whether it actually does occur is a quantitative question concerning the strength of the induced ferromagnetic interaction. Thus, such an entropy-driven phase transition is believed not to occur in the 3-state Potts antiferromagnet on the square lattice \(\mathbb{Z}^2\); but Monte Carlo evidence \([71, 31, 32]\) suggests that it does occur in this same model on the simple-cubic lattice \(\mathbb{Z}^3\) and presumably also on \(\mathbb{Z}^d\) for all \(d \geq 3\); moreover, Peled \([62]\) and Galvin et al. \([27]\) have recently proven this for all sufficiently large \(d\) and also for a “thickened” version of \(\mathbb{Z}^2\) \([62]\).

From the point of view of the Peierls argument, the relevant issue is the strength of the entropic penalty for domain walls between differently-ordered regions, compared to the entropy associated to those domain walls. In order to successfully carry out the Peierls argument, one must consider all the relevant ordered phases, find an appropriate definition of Peierls contours separating spatial regions resembling those

---

\(^9\) Some exceptions discussed in the physics literature prior to \([47]\) were the constrained square-lattice 4-state antiferromagnetic Potts model \([11]\) and the triangular-lattice antiferromagnetic spin-\(s\) Ising model for large enough \(s\) \([73]\), both of which appear to lie in a non-critical ordered phase at zero temperature.

\(^10\) It can also occur in antiferromagnetic models on non-bipartite lattices: for instance, in the 4-state Potts antiferromagnets on the union-jack and bisected-hexagonal lattices \([14]\), which are tripartite, and for which ordering on one sublattice increases the entropy available to the other two sublattices. However, we are concerned here for simplicity with the bipartite case.
ordered phases, and prove that long Peierls contours $\gamma$ are suppressed like $e^{-c|\gamma|}$ with a sufficiently large constant $c$.

The simplest situation arises when there is an asymmetry between the two sublattices, so that it is entropically more favorable for one of them (say, $V_0$) to be ferromagnetically ordered and for the other ($V_1$) to carry all the entropy. This situation is expected to occur, for instance, if $V_1$ has a higher density of points than $V_0$. The case treated in this paper, in which $G_0$ is a triangulation, achieves this in the strongest possible way: namely, for the Euclidean lattices in our class, it is easy to see using Euler’s formula that the spatial densities of the sublattices $V_0$ and $V_1$ are in the proportion $1:2$, which is the most extreme ratio achievable for two dual periodic Euclidean lattices.

In this asymmetric situation, one knows in advance which sublattice ($V_0$) is going to be ferromagnetically ordered (if the entropic effect is strong enough to produce any long-range order at all); therefore, for the 3-state Potts antiferromagnet on $G$, one expects at low temperature to have (at least) three distinct ordered phases, corresponding to the three possible choices for the color that dominates on $V_0$. One may therefore define Peierls contours just as one would for a ferromagnetic Potts model on $G_0$ (see Section 2.1 below for details), and then try to show that long Peierls contours are sufficiently suppressed, i.e., that it is sufficiently costly to create an interface between regions where one and another color are used on $V_0$. This is a quantitative problem, which is made difficult by the fact that (unlike in a ferromagnetic model) one does not have any parameter that can be varied to make the suppression of long contours as large as one wishes.

The situation is even more delicate for lattices, such as $\mathbb{Z}^d$, where the two sublattices play a symmetric role (in the sense that there exists an automorphism of $G$ carrying one sublattice onto the other). Indeed, for models with symmetry between the sublattices, for every Gibbs measure where one sublattice is ordered (in the sense of being colored more often with one preferred color), there must obviously exist a corresponding Gibbs measure where the other sublattice is ordered. Therefore, the system has two “choices” to make: first, of the sublattice to be ordered, and then of the color in which it is ordered — which leads to a total of (at least) six distinct ordered phases for the 3-state model. For this sort of long-range order to occur, it must be sufficiently costly to create an interface between any pair of distinct ordered phases; in particular, it must be costly to create an interface between regions where one and the other sublattice are ordered (in whatever colors). To prove such a result will almost certainly require a different (and more subtle) definition of Peierls contour than is used in the asymmetric case.

The example of $\mathbb{Z}^d$ for $d$ large [27] shows that asymmetry is not necessary for the existence of a finite-temperature phase transition. But one can nevertheless say...

---

11 We say that a subset $W \subseteq V$ has spatial density $\lambda$ if, for any sequence of finite sets $V_{(n)}$ increasing to $V$ such that proportion of vertices in $V_{(n)}$ that are adjacent to $V \setminus V_{(n)}$ tends to zero (i.e., van Hove–Følner convergence), the fraction $|V_{(n)} \cap W|/|V_{(n)}|$ tends to $\lambda$. 
heuristically that asymmetry enhances the effect driving the transition, by increasing the strength of the effective ferromagnetic interaction on the favored sublattice (while of course decreasing it on the disfavored sublattice).

Everything said so far holds for an arbitrary bipartite lattice. But the case in which $G$ is a plane quadrangulation is special, because $G_0$ and $G_1$ are not merely the two sublattices: they are a dual pair of plane graphs. In particular, there is a symmetry between $G_0$ and $G_1$ if and only if $G_0$ is self-dual; and there may be special reasons, connected with the topology of the plane, that make this self-dual case special (e.g. critical at zero temperature). Now, it is well known that the square lattice is self-dual; what seems to be less well known is that there exist many other examples of self-dual periodic plane graphs, including the “hextri” lattice [60, Fig. 1] [67, Fig. 16], the “house” lattice [60, Fig. 2], and the martini-B lattice [66, Fig. 8]. Preliminary results of Monte Carlo simulations on a variety of plane quadrangulations suggest that

(a) If $G_0$ is self-dual, then the 3-state Potts antiferromagnet on the associated quadrangulation $G$ has a zero-temperature critical point; and

(b) If $G_0$ is not self-dual, then the 3-state Potts antiferromagnet on $G$ has (always? usually?) a finite-temperature phase transition.

In other words, it seems that for plane quadrangulations — unlike for general bipartite lattices — asymmetry may be both necessary and sufficient for the existence of a finite-temperature phase transition. It would be very interesting to find a deeper theoretical explanation, and ultimately a proof, of this apparent fact. We conjecture that there is an exact duality mapping that explains why (a) is true. As for (b), one could argue for it heuristically as follows: Because the model at zero temperature has a height representation, it should be either critical or ordered. If the self-dual cases are critical, then the non-self-dual cases should be ordered, since asymmetry enhances the phase transition; and if the self-dual cases are ordered, then the non-self-dual cases should be even more strongly ordered. It goes without saying that this heuristic argument is extremely vague — no criterion for comparing lattices is given — and hence very far from suggesting a strategy of proof.

Entropy-driven phase transitions are also possible in the $q$-state Potts antiferromagnet for $q > 3$, but now one must consider the possibility of Gibbs measures associated to other partitions $[q] = Q_0 \cup Q_1$, in which the vertices in $V_0$ (resp. $V_1$) take predominantly colors from $Q_0$ (resp. $Q_1$). Depending on the size and shape of $V_0$ and $V_1$ and the value of $q$, such measures might be entropically favored. For instance, such ordering with $|Q_0| = |Q_1| = 2$ has been claimed to occur in the 4-state

\[\text{(12)}\] For any connected bipartite graph $G = (V,E)$ with vertex bipartition $V = V_0 \cup V_1$, one can define graphs $G_0 = (V_0,E_0)$ and $G_1 = (V_1,E_1)$ by setting $E_0 = \{\{u,v\}: u,v \in V_0 \text{ and } d_G(u,v) = 2\}$ and likewise for $E_1$. But if $G$ is non-planar, or is planar but not a quadrangulation, it is not clear whether these definitions will be useful.

\[\text{(13)}\] Including to the authors until very recently.
Potts antiferromagnet on the simple-cubic lattice $\mathbb{Z}^3$. Naive entropic considerations suggest that if the densities of the sublattices $V_0$ and $V_1$ are in the ratio $\alpha:1-\alpha$, then the dominant ordering would have $|Q_0| \approx \alpha q$. In general, one would expect to have $\binom{q}{|Q_0|}$ ordered phases in the asymmetric case, and $2^\lfloor\frac{q}{|Q_0|}\rfloor = 2^\lfloor\frac{q}{\lfloor q/2 \rfloor}\rfloor$ in the symmetric case. The cases with $|Q_0| > 1$ will require a different (and more subtle) definition of Peierls contour than the one used here for $|Q_0| = 1$.

The foregoing considerations are purely entropic; a more complicated phase diagram, involving tradeoffs between entropy and energy, can presumably be obtained by adding additional couplings into the Hamiltonian (1.4). Suppose, for instance, that in the 3-state Potts antiferromagnet on a plane quadrangulation $G$ where $G_0$ is a triangulation, we add an explicit ferromagnetic interaction, of strength $\lambda$, between adjacent vertices in the sublattice $G_1$. Then for small $\lambda$ we expect that the favored ordering at low temperature will be the same as for $\lambda = 0$, namely monocolor on $V_0$ and bicolor on $V_1$; but for large positive $\lambda$ the favored ordering will instead be monocolor on $V_1$ and bicolor on $V_0$. It is then natural to guess that for large $\beta$ there is either a switchover between the two orderings at some particular value $\lambda_t(\beta)$, or else a pair of phase transitions $\lambda_1(\beta) < \lambda_2(\beta)$ with a disordered phase in-between [and possibly $\lambda_1(\infty) = \lambda_2(\infty)$]. It is an interesting open question to determine the correct qualitative phase diagram in the $(\lambda, \beta)$-plane and the order of the phase transition(s).

1.4. Some further open problems. Here are some further open problems suggested by our work:

1) Prove (or disprove) that

(a) the finite-volume measures $\mu_{k,\beta}^\Lambda$ used in the proof of Theorem 1.1 (see Section 2.3 below) converge as $\Lambda \uparrow V$ (i.e., there is no need to take a subsequence);
(b) the resulting infinite-volume Gibbs measures $\mu_{\beta,k}$ are extremal Gibbs measures; and
(c) $\mu_{\beta,k}$ are invariant with respect to the automorphism of the graph $G$.

2) Prove (or disprove) that for our lattices there are no more than three extremal translation-invariant Gibbs measures at small but strictly positive temperature. For this, one would need to control more general boundary conditions than the uniform colorings on $V_0$ that we have used here.

Please note that at zero temperature, there are in fact more than three extremal infinite-volume Gibbs measures on the diced lattice, since there exist ground-state configurations $\tau$, similar to the example on $\mathbb{Z}^2$ sketched in Remark 2 after Theorem 1.1, such that for any finite $\Lambda \subset V$ there exists only one ground state that agrees with $\tau$ on $V \setminus \Lambda$ (namely, $\tau$ itself). The delta measure on such a ground state is therefore a zero-temperature Gibbs measure; but by the argument sketched at
the end of that Remark, this Gibbs measure is not a limit of positive-temperature Gibbs measures.

It is worth pointing out, however, that this latter argument makes essential use of the fact that the lattice is Euclidean (in particular, its isoperimetric constant is zero). This raises the question whether on hyperbolic lattices there might exist delta-measure zero-temperature Gibbs measures that are limits of positive-temperature Gibbs measures.

3) Extend these techniques to the $q$-state Potts antiferromagnet with $q > 3$ on suitable lattices. For instance, one might hope to prove the existence of an entropy-driven phase transition in the $q$-state Potts antiferromagnet on $\mathbb{Z}^d$ for suitable pairs $(q, d)$, i.e., for $q < \min\{q_c(\mathbb{Z}^d)\}$. In this case it is not completely clear, even heuristically, how $q_c(\mathbb{Z}^d)$ should behave as $d \to \infty$. The example of the infinite $\Delta$-regular tree, which has multiple Gibbs measures when $q \leq \Delta$ [8] and a unique Gibbs measure when $q \geq \Delta + 1$ [42], suggests that we might have $q_c(\mathbb{Z}^d) \approx 2d$.

1.5. Plan of this paper. The remainder of this paper is organized as follows: In Section 2 we introduce the Peierls-contour representation of our model and sketch the main ideas underlying our proofs. In particular, we formulate the key steps in our proof as precise lemmas (Lemmas 2.1–2.5) that will be proven later, and we show how they together imply Theorems 1.1 and 1.2. In Section 3 we prove Lemmas 2.1 and 2.5 in the zero-temperature case $\beta = \infty$, using a Peierls argument. In Section 4 we extend these proofs to the low-temperature case $\beta \geq \beta_0$, and we also prove the technical Lemmas 2.2 and 2.4. In Section 5 we use a random-cluster argument to deduce positive magnetization (Lemma 2.3). In the Appendix we review the needed theory of infinite graphs.

2. Main structure of the proofs

2.1. Contour model. Our proofs of Theorems 1.1 and 1.2 are based on suitable bounds for finite-volume Gibbs measures, uniform in the system size and in the inverse temperature above a certain value. We will concentrate on finite-volume Gibbs measures with uniform 1 boundary conditions on the sublattice $V_0$; by symmetry, all statements immediately imply analogous results for boundary conditions 2 or 3. We will always employ finite sets $\Lambda \subset V$ whose external boundary lies entirely in the sublattice $V_0$, i.e. $\partial \Lambda \subset V_0$. We will also assume that $\Lambda \subset V$ is simply connected, by which we mean that both $\Lambda$ and $V \setminus \Lambda$ are connected in $G$. Thus, let us fix any configuration $\tau$ that equals 1 on $V_0$, and let $\mu_{\Lambda, \beta}^1$ denote the finite-volume Gibbs measure in $\Lambda$ with boundary condition $\tau$ and inverse temperature $\beta \in [0, \infty)$. (Since $\partial \Lambda \subset V_0$, this measure is the same for all configurations $\tau$ that equal 1 on $V_0$.)

Our proofs are based on a version of Peierls argument relying on a contour reformulation of the measure $\mu_{\Lambda, \beta}^1$. Our goal is to prove that the sublattice $V_0$ exhibits ferromagnetic order of a suitable kind. Therefore we will define Peierls contours just as one would for studying the ferromagnetic Potts model on $G_0$. Thus, for any
configuration $\sigma$, we look only at the restriction of $\sigma$ to $V_0$, and we define $E_0(\sigma)$ to be the set of “unsatisfied edges”, i.e.

$$E_0(\sigma) := \{ \{u, v\} \in E_0: u, v \in V_0 \cap (\Lambda \cup \partial \Lambda) \text{ and } \sigma_u \neq \sigma_v \}. \quad (2.1)$$

Letting

$$E_1(\sigma) := \{ e \in E_1: e \text{ crosses some } f \in E_0(\sigma) \} \quad (2.2)$$

denote the edges in the dual graph $G_1$ that cross an edge in $E_0(\sigma)$, we see that edges in $E_1(\sigma)$ correspond to boundaries separating areas where the vertices of $G_0$ are uniformly colored in one of the colors 1, 2, 3. Note that since $\partial \Lambda \subset V_0$ and $\partial \Lambda$ is uniformly colored, every edge of $E_1(\sigma)$ has both of its endpoints in $\Lambda$.

Since $G_0$ is a triangulation, each vertex of $G_1$ is of degree 3. If the three vertices of $G_0$ surrounding a vertex $v \in V_1$ are colored with three different colors, then one of these vertices must have the same color as $v$. This is clearly not possible for a ground state $\sigma$ (i.e., a proper coloring), so at $\beta = \infty$ at most two different colors can surround any vertex $v \in V_1$. It follows that at zero temperature, either zero or two edges of $E_1(\sigma)$ emanate from the vertex $v$. Hence $E_1(\sigma)$ consists of a collection $\Gamma(\sigma)$ of disjoint simple circuits that we call contours.

At positive temperature, we define contours to be connected components of $E_1(\sigma)$, which can be much more complicated than a circuit. Nevertheless, we will show that at low temperatures, contours that are not simple circuits are rare.

2.2. The basic lemmas. Let us now sketch in broad lines the main ideas of our proofs, and formulate a number of precise lemmas, to be proven later, that together will imply our main results. We have seen that uniformly colored areas in the sublattice $G_0$ are separated by contours in the sublattice $G_1$, which at zero temperature are simple circuits. The number of different simple circuits of a given length $L$ surrounding a given point is roughly of order $\alpha^L$, where $\alpha$ is the connective constant of the lattice $G_1$. Since each vertex in $G_1$ has degree 3, a contour entering a vertex has two possible directions in which to continue. In view of this, it is easy to see that $\alpha \leq 2$. With a bit more work using quasi-transitivity, this can be improved to $\alpha < 2$. On the other hand, each vertex $v$ in $G_1$ that lies on a contour is surrounded by vertices in $G_0$ of two different colors. At zero temperature, this means that there is only one color available for $v$, compared to two for a vertex in $G_1$ that does not lie on a contour. As a result, for each contour of length $L$ we have to pay an entropic price $2^{-L}$. In view of this, we will prove in Section 3 below that the expected number of contours surrounding a given site is of order $\sum_L \alpha^L 2^{-L}$, which is finite.

Note that this reasoning tells us that the Peierls sum is finite, but not necessarily that it is small. In a traditional Peierls argument (such as, for example, the proof of [51, Theorem IV.3.14]), one argues that if the Peierls contour sum is smaller than a certain model-dependent threshold (typically a number somewhat less than 1), then the model has spontaneous magnetization. This is indeed how we will prove Theorem 1.2 for the diced lattice. But for the general class of lattices in Theorem 1.1, all one can hope to prove is that the Peierls sum is finite; it need not be small. To
handle this situation, we use a trick that we learned from [22, section 6a], where it is used for percolation. We observe that if $\Delta_0 \subset V_0$ is connected in $G_0$, then $\Delta_0$ is uniformly colored in one color if and only if no contours cut through $\Delta_0$. On the other hand, if $\Delta_0$ is sufficiently large, then by the finiteness of the Peierls sum, $\Delta_0$ is unlikely to be surrounded by a contour. It follows that, conditional on $\Delta_0$ being uniformly colored (which is of course a rare event), it is much more likely for $\Delta_0$ to be uniformly colored in the color 1 than in either of the other two colors.

More precisely, for each $k \in \{1, 2, 3\}$ and each finite set $\Delta_0 \subset V_0$, let $J_{k,\Delta_0}$ denote the event that all sites in $\Delta_0$ have the color $k$, and let $J_{\Delta_0} = \bigcup_{k=1}^{3} J_{k,\Delta_0}$ denote the event that all sites in $\Delta_0$ are colored with the same color. By the arguments sketched above for the case of zero temperature, together with a careful estimate of non-simple contours at small positive temperatures, we are able to prove the following lemma:

**Lemma 2.1. (Long-range dependence)** There exists $\beta_0 < \infty$ such that for each $\epsilon > 0$, there exists $M_\epsilon < \infty$ such that for every finite set $\Delta_0 \subset V_0$ that is connected in $G_0$ and satisfies $|\Delta_0| \geq M_\epsilon$, one has

$$
\mu_{\Lambda,\beta}^1(\mathcal{J}, \Delta_0 \mid \mathcal{J}) \geq 1 - \epsilon
$$

uniformly for all $\beta \in [\beta_0, \infty]$ and all simply connected finite sets $\Lambda \supseteq \Delta_0$ such that $\partial \Lambda \subset V_0$.

In order for Lemma 2.1 to be of any use, we need to show that the event on which we are conditioning in (2.3) has positive probability, uniformly in the system size:

**Lemma 2.2. (Uniformly colored sets)** Let $\Delta_0 \subset V_0$ be finite and connected in $G_0$. Then there exists a constant $\delta > 0$ such that

$$
\mu_{\Lambda,\beta}^1(\mathcal{J}, \Delta_0) \geq \delta
$$

uniformly for all $0 \leq \beta \leq \infty$ and all finite and simply connected $\Lambda \supseteq \Delta_0$ such that $\partial \Lambda \subset V_0$.

Of course $\delta$ gets very small as $\Delta_0$ gets large, but we do not care, as we will take $\Delta_0$ to be large but fixed.

Let us note that Lemmas 2.1 and 2.2 are sufficient, by themselves, to prove the existence of at least three distinct infinite-volume Gibbs measures at all $\beta \in [\beta_0, \infty]$. These infinite-volume Gibbs measures may or may not have spontaneous magnetization, but they do at least have long-range order of a special kind: namely, they assign unequal probabilities to the (rare) events $\mathcal{J}_{k,\Delta_0}$ ($k = 1, 2, 3$) for some large but finite set $\Delta_0$. This part of the argument is quite general and applies to other

---

14 To see this, just follow the proof of Theorem 1.1 given in Section 2.3 below and disregard all references to Lemma 2.3. The bound (2.10) and its analogues for $k = 2, 3$ survive to the (subsequential) infinite-volume limit and hence show that the Gibbs measures $\mu_{k,\beta}$ for $k = 1, 2, 3$ are distinct.
models as well, as long as it can be shown that the Peierls sum is finite, even if it is not necessarily small.\textsuperscript{[15]}

But for our particular model, we can actually do better and prove that there is spontaneous magnetization, thanks to the following lemma, which says that if a sufficiently “thick” block is more likely to be uniformly colored in one color than in the other two colors, then the same must be true for single sites within that block. Let us say that a set \( \Delta \subset V \) is thick\textsuperscript{[16]} if there exists a nonempty finite subset \( \Delta_1 \subset V_1 \) that is connected in \( G_1 \) and such that \( \Delta = \{ v \in V : d_G(v, \Delta_1) \leq 1 \} \). Then \( \Delta \) is connected in \( G \), and we have \( \Delta_1 = \Delta \cap V_1 \); we write \( \Delta_0 := \Delta \cap V_0 \).

**Lemma 2.3. (Positive magnetization)** Fix \( \beta_0 > 0 \) and let \( \Delta \subset V \) be thick. Then there exists \( \epsilon > 0 \) such that for each \( v_0 \in \Delta_0 \),

\[
(2.5) \quad \mu^1_{\Lambda, \beta}(\sigma_{v_0} = 1) - \mu^1_{\Lambda, \beta}(\sigma_{v_0} = 2) \geq \epsilon \left[ \mu^1_{\Lambda, \beta}(\mathcal{J}_1, \Delta_0) - \mu^1_{\Lambda, \beta}(\mathcal{J}_2, \Delta_0) \right],
\]

and similarly, for each \( v_1 \in \Delta_1 \),

\[
(2.6) \quad \mu^1_{\Lambda, \beta}(\sigma_{v_1} = 2) - \mu^1_{\Lambda, \beta}(\sigma_{v_1} = 1) \geq \epsilon \left[ \mu^1_{\Lambda, \beta}(\mathcal{J}_1, \Delta_0) - \mu^1_{\Lambda, \beta}(\mathcal{J}_2, \Delta_0) \right],
\]

uniformly for all \( \beta \in [\beta_0, \infty] \) and all simply connected finite sets \( \Lambda \supseteq \Delta \) such that \( \partial \Lambda \subset V_0 \).

The proof of Lemma 2.3 is not very complicated but is very much dependent on the specific properties of our Potts model. Inspired by the cluster algorithm introduced in [70, 71], we condition on the position of the 3’s and use the random-cluster representation for the Ising model of 1’s and 2’s on the remaining diluted lattice. We show that the difference of probabilities that \( \Delta_0 \) is uniformly colored in the color 1 or in the color 2 equals the probability that \( \Delta_0 \) is uniformly colored and there is a 1-2 random-cluster connection between \( \Delta_0 \) and the boundary of \( \Lambda \). Using this latter quantity, it is then easy to produce (by a finite-energy argument) a lower bound on the probability that there is a 1-2 random-cluster connection between a fixed lattice site \( v_0 \) and the boundary of \( \Lambda \); and this, in turn, equals the magnetization.

The last main missing ingredient of Theorem 1.1 is the following lemma, which shows that improperly colored edges are rare when \( \beta \) is large; in particular it shows that any limit as \( \beta \to \infty \) of the finite-temperature infinite-volume Gibbs measures that we will construct is concentrated on the set \( S_g \) of ground states.

\textsuperscript{[15]} For example, consider the Ising model on \( \mathbb{Z}^2 \), with the contours constructed as in [30] Lemma 6.14] by taking a circuit on the dual lattice that is the external boundary of the connected (by nearest-neighbor edges in \( \mathbb{Z}^2 \)) region of constant spin containing origin. Then this argument gives the existence of at least two distinct infinite-volume Gibbs measures whenever \( e^{-2\beta} \alpha_{SQ} < 1 \) (i.e. \( \beta > \frac{1}{2} \log \alpha_{SQ} \)), where \( \alpha_{SQ} \) is the connective constant for self-avoiding polygons (or equivalently, self-avoiding walks [33 Corollary 3.2.5]) on the square lattice, and \( \beta \) is the inverse temperature in the standard Ising normalization. (It is known that \( \alpha_{SQ} < 2.679192495 \) [63]; the current best numerical estimate is \( \alpha_{SQ} \approx 2.63815853031(3) \) [40].

\textsuperscript{[16]} Notice that the sublattices \( V_0 \) and \( V_1 \) enter the definition of thickness in an asymmetric way.
Lemma 2.4. (Rarity of improperly colored edges) There exists $C < \infty$ such that

\begin{equation}
\mu^1_{\Lambda,\beta}(\sigma_u = \sigma_v) \leq Ce^{-\beta}
\end{equation}

for all $\beta \in [0, \infty]$, all $\{u,v\} \in E$, and all finite and simply connected $\Lambda \ni u,v$ such that $\partial \Lambda \subset V_0$.

Finally, to prove Theorem 1.2 for the diced lattice, we need the following quantitative bound:

Lemma 2.5. (Explicit Peierls bound for the diced lattice) If $G$ is the diced lattice, then there exists $C < \infty$ such that

\begin{equation}
\mu^1_{\Lambda,\beta}(\sigma_{v_0} = 1) \geq 0.90301 - Ce^{-\beta}
\end{equation}

uniformly for all $\beta \in [0, \infty]$, all $v_0 \in V_0$, and all simply connected finite sets $\Lambda \ni v_0$ such that $\partial \Lambda \subset V_0$.

2.3. Proof of the main theorems, given the basic lemmas. Let us now show how to prove Theorems 1.1 and 1.2 given Lemmas 2.1–2.5.

Proof of Theorem 1.1. Fix $\epsilon > 0$, let $\beta_0, M_\epsilon$ and $\Delta_0$ be as in Lemma 2.1, and let $\delta$ be as in Lemma 2.2. Since the colors 2 and 3 play a symmetric role under the measure $\mu^1_{\Lambda,\beta}$, we have

\begin{equation}
\mu^1_{\Lambda,\beta}(J_{2,\Delta_0} | J_{\Delta_0}) = \frac{1}{2} \left[ 1 - \mu^1_{\Lambda,\beta}(J_{1,\Delta_0} | J_{\Delta_0}) \right]
\end{equation}

and hence

\begin{align}
\mu^1_{\Lambda,\beta}(J_{1,\Delta_0}) - \mu^1_{\Lambda,\beta}(J_{2,\Delta_0}) &= \frac{1}{2} \left[ 3 \mu^1_{\Lambda,\beta}(J_{1,\Delta_0} | J_{\Delta_0}) - 1 \right] \mu^1_{\Lambda,\beta}(J_{\Delta_0}) \\
&\geq \frac{1}{2} [3(1 - \epsilon) - 1] \delta = \frac{1}{2} (2 - 3\epsilon) \delta,
\end{align}

which is positive for $\epsilon < 2/3$ (which we henceforth assume). Then, for any $v_0 \in V_0$, we may choose a thick set $\Delta \subset V$ such that $|\Delta_0| \geq M_\epsilon$ and $v_0 \in \Delta_0$ (with $\Delta_0 := \Delta \cap V_0$ as defined earlier). By (2.10) together with Lemma 2.3, there exists $\bar{\epsilon}(v_0) > 0$ such that

\begin{equation}
\mu^1_{\Lambda,\beta}(\sigma_{v_0} = 1) - \mu^1_{\Lambda,\beta}(\sigma_{v_0} = 2) \geq \bar{\epsilon}(v_0)
\end{equation}

uniformly for all $\beta \in [\beta_0, \infty]$ and all finite and simply connected $\Lambda \ni \Delta_0$ such that $\partial \Lambda \subset V_0$. Since the measure $\mu^1_{\Lambda,\beta}$ treats the colors 2 and 3 symmetrically, it follows that

\begin{equation}
\mu^1_{\Lambda,\beta}(\sigma_{v_0} = 1) - \frac{1}{2} \left[ 1 - \mu^1_{\Lambda,\beta}(\sigma_{v_0} = 1) \right] \geq \bar{\epsilon}(v_0)
\end{equation}

and hence

\begin{equation}
\mu^1_{\Lambda,\beta}(\sigma_{v_0} = 1) \geq \frac{1}{3} + \frac{2}{3} \bar{\epsilon}(v_0).
\end{equation}
Similarly, for any \( v_1 \in V_1 \), we may choose a thick set \( \Delta \) such that \( |\Delta_0| \geq M_\epsilon \) and \( v_1 \in \Delta_1 \). An analogous argument then shows that

\[
\mu_{\Lambda,\beta}^1(\sigma_{v_1} = 1) \leq \frac{1}{3} - \frac{2}{3}\tau(v_1),
\]

again uniformly in \( \Lambda \).

In this argument \( \tau \) depends on \( v_0 \) or \( v_1 \). But since all the quantities under study are invariant under graph automorphisms, \( \tau \) actually depends only on the type of \( v_0 \) or \( v_1 \). And since by quasi-transitivity there are only finitely many types, we may choose \( \epsilon \) such that (2.13) and (2.14) hold uniformly for all \( v_0 \in V_0 \) and \( v_1 \in V_1 \).

To construct the desired infinite-volume Gibbs measures, we use a compactness argument. For any \( \beta \in [0, \infty) \) and finite \( \Lambda \subset V \), let \( \overline{\mu}_{\Lambda,\beta} : = \mu_{\Lambda,\beta}^1 \otimes \delta_{V \setminus \Lambda} \), i.e., if \( \sigma \in \{1,2,3\}^V \) is distributed according to \( \overline{\mu}_{\Lambda,\beta} \), then \( (\sigma_v)_{v \in \Lambda} \) is distributed according to \( \mu_{\Lambda,\beta}^1 \) and \( \sigma_v = \tau_v \) for all \( v \in V \setminus \Lambda \). Choose finite and simply connected \( \Lambda_n \uparrow V \) such that \( \partial \Lambda_n \subset V_0 \). Since \( S = \{1,2,3\}^V \) is a compact space, the set of measures \( \{\overline{\mu}_{\Lambda_n,\beta}\} \) is automatically tight. It follows from [30, Theorem 4.17] that each weak subsequential limit \( \mu_\beta \) as \( \Lambda_n \uparrow \Lambda \) is an infinite-volume Gibbs measure at inverse temperature \( \beta \). Taking the limit \( \Lambda_n \uparrow V \) in (2.13)/(2.14), we see that

\[
\mu_\beta(\sigma_{v_0} = 1) \geq \frac{1}{3} + \epsilon \quad \text{for} \quad v_0 \in V_0
\]

\[
\mu_\beta(\sigma_{v_1} = 1) \leq \frac{1}{3} - \epsilon \quad \text{for} \quad v_1 \in V_1
\]

Taking the limit \( \Lambda_n \uparrow V \) in Lemma 2.4, we obtain

\[
\mu_\beta(\sigma_u = \sigma_v) \leq Ce^{-\beta} \quad \text{for} \quad \{u,v\} \in E.
\]

\[\square\]

**Proof of Theorem 1.2.** (a) and (c) follow from the same arguments as in the proof of Theorem 1.1, but with the inequality (2.13) replaced by (2.8).

To prove (b), consider any \( v_1 \in V_1 \cap \Lambda \) and let \( w_1, w_2, w_3 \in V_0 \cap (\Lambda \cup \partial \Lambda) \) be its neighbors in \( G \). Then the DLR equations for the volume \( \{v_1\} \) imply that

\[
\mu_{\Lambda,\beta}^1(\sigma_{v_1} = 1|\sigma_{w_1} = \sigma_{w_2} = \sigma_{w_3} = 1) = \frac{e^{-3\beta}}{2 + e^{-3\beta}} \quad (2.18a)
\]

\[
\mu_{\Lambda,\beta}^1(\sigma_{v_1} = 1|\sigma_{w_1} = \sigma_{w_2} = 1, \sigma_{w_3} \neq 1) = \frac{e^{-2\beta}}{1 + e^{-\beta} + e^{-2\beta}} \quad (2.18b)
\]

(we call these the “good” cases). In the “bad” cases (i.e., those with two or three spins \( \sigma_{w_i} \neq 1 \)) we will use only that the conditional probability is \( \leq 1 \). On the other hand, using (a) we can bound the probability of the “bad” cases by

\[
\mu_{\Lambda,\beta}^1(\text{two or three spins} \sigma_{w_i} \neq 1) \leq \frac{1}{2} \mathbb{E}_{\mu_{\Lambda,\beta}^1}(\# \text{ spins} \sigma_{w_i} \neq 1)
\]

\[
\leq \frac{1}{2} \times 3 \times (1 - 0.90301 + Ce^{-\beta})
\]

\[
(2.19) \quad \leq 0.14549 + C'e^{-\beta},
\]

\[\text{LOW-TEMPERATURE POTTS ANTIFERROMAGNETS} \quad 19\]
and we bound the probability of the “good” cases trivially by 1. Putting together (2.18a,b) and (2.19), we conclude that 
\[ \mu_{\Lambda,\beta}(\sigma_v = 1) \leq 0.14549 + C'e^{-\beta} + e^{-2\beta}. \]

\[ \square \]

3. The zero-temperature case

In the present section, we will prove Lemmas 2.1 and 2.5 in the zero-temperature case \( \beta = \infty \). Then, in Section 4 we will show how our arguments can be adapted to cover the (more complicated) case of low positive temperatures; there we will also prove the technical Lemmas 2.2 and 2.4. The proof of Lemma 2.3 is postponed to Section 5.

3.1. Contour model for zero temperature. Let \( \Lambda \subset V \) be finite and simply connected in \( G \) and such that \( \partial \Lambda \subset V_0 \). Recall that \( \mu_{\Lambda,\infty} \) is the uniform distribution on the set \( (S_g)_{\Lambda}^1 \) of all proper 3-colorings of \( \Lambda \cup \partial \Lambda \) that take the color 1 on \( \partial \Lambda \), i.e. all configurations \( \sigma \in \{1,2,3\}^{\Lambda \cup \partial \Lambda} \) such that \( \sigma_u \neq \sigma_v \) for all \( u, v \in \Lambda \cup \partial \Lambda \) with \( \{u,v\} \in E \) and \( \sigma_v = 1 \) for all \( v \in \partial \Lambda \). Since \( \partial \Lambda \subset V_0 \), the set \( (S_g)_{\Lambda}^1 \) is nonempty: for instance, it includes all configurations in which all sites of \( V_0 \) are colored 1 and all sites of \( V_1 \) are colored 2 or 3.

As explained in Section 2.1, for any \( \sigma \in (S_g)_{\Lambda}^1 \), we let \( E_1(\sigma) \) be the collection of edges in \( G_1 \) that separate areas where the vertices of \( G_0 \) are uniformly colored in one of the colors 1, 2, 3. And since at zero temperature at most two different colors on \( V_0 \) can meet at any vertex in \( V_1 \), the set \( E_1(\sigma) \) consists of a collection \( \Gamma(\sigma) \) of disjoint simple circuits that we call contours. [This is what makes the zero-temperature case so easy to handle. At positive temperature, a connected component of \( E_1(\sigma) \) can be much more complicated than a circuit: see Section 4 below.]

In the zero-temperature case, therefore, we use the term contour to denote any simple circuit in \( G_1 \). We write \( |\gamma| \) to denote the length of a contour \( \gamma \), defined as the number of its edges (or equivalently the number of its vertices). For a collection \( \Gamma \) of disjoint contours, we write \( |\Gamma| := \sum_{\gamma \in \Gamma} |\gamma| \) for the total length of the contours in \( \Gamma \), and \( \#\Gamma \) for the number of contours in \( \Gamma \). Each contour \( \gamma \) divides \( V_0 \) into two connected (in the sense of \( G_0 \) components, of which one is infinite and the other is finite and simply connected. We call these the exterior \( \text{Ext}(\gamma) \) and interior \( \text{Int}(\gamma) \) of \( \gamma \), respectively. We will say that a contour \( \gamma \) surrounds \( \Delta_0 \) if \( \Delta_0 \subseteq \text{Int}(\gamma) \). We say that a contour lies in \( \Lambda \) if all its vertices are in \( V_1 \cap \Lambda \). Note that if \( \gamma \) lies in \( \Lambda \), then by our assumption that \( \Lambda \) is simply connected, we have \( \text{Int}(\gamma) \subseteq \Lambda \).

To each configuration \( \sigma \in (S_g)_{\Lambda}^1 \), there thus corresponds a unique collection \( \Gamma(\sigma) \) of disjoint contours in \( \Lambda \). Conversely, to each collection \( \Gamma \) of disjoint contours, there are \( 2^{\#\Gamma} 2^{|V_1 \cap \Lambda| - |\Gamma|} \) distinct configurations \( \sigma \in (S_g)_{\Lambda}^1 \) that yield the collection \( \Gamma(\sigma) = \Gamma \). Here the first and second factor are the number of restrictions \( \sigma_{V_0 \cap \Lambda} \) and \( \sigma_{V_1 \cap \Lambda} \), respectively, that are consistent with the specified collection of contours and the fixed boundary condition \( \sigma_v = 1 \) for \( v \in \partial \Lambda \). To understand the first factor, observe that in passing through any contour (from outside to inside) we have 2
(= q – 1) independent alternatives for the choice of the color on \( V_0 \) just inside the contour. The second factor comes from the fact that there are either one (= q – 2) or two (= q – 1) colors available for a vertex in \( G_1 \), depending on whether this vertex lies on a contour or not. Notice that, given \( \Gamma \), this latter number is independent of the configuration \( \sigma_{V_0 \cap \Lambda} \).

Let us therefore introduce the probability measure

\[
\nu_{\Lambda}(\Gamma) = \frac{1}{Z_{\Lambda}} 2^{\# \Gamma} 2^{-|\Gamma|}
\]

on the set of all collections \( \Gamma \) of disjoint contours in \( \Lambda \), where \( Z_{\Lambda} = \sum_{\Gamma} 2^{\# \Gamma} 2^{-|\Gamma|} \) is the normalizing constant. We have just shown that, under the probability measure \( \mu_{1,\Lambda,\infty} \), the contour configuration \( \Gamma(\sigma) \) is distributed according to \( \nu_{\Lambda} \).

The first step in any Peierls argument is to obtain an upper bound on the probability that \( \Gamma \) contains a given contour \( \gamma \):

**Lemma 3.1. (Bound on probability of a contour)** Let \( \Lambda \subset V \) be a simply connected finite set such that \( \partial \Lambda \subset V_0 \). Then, for any contour \( \gamma \) in \( \Lambda \),

\[
\nu_{\Lambda}(\{ \Gamma : \gamma \in \Gamma \}) \leq \frac{2^{1-|\gamma|}}{1 + 2^{1-|\gamma|}}.
\]

**Proof.** We have

\[
\nu_{\Lambda}(\{ \Gamma : \gamma \in \Gamma \}) = \sum_{\Gamma \ni \gamma} \nu_{\Lambda}(\Gamma) = 2^{1-|\gamma|} \sum_{\Gamma \ni \gamma} \nu_{\Lambda}(\Gamma \setminus \{\gamma\})
\]

\[
\leq 2^{1-|\gamma|} \sum_{\Gamma \ni \gamma} \nu_{\Lambda}(\Gamma) = 2^{1-|\gamma|} \left[ 1 - \nu_{\Lambda}(\{ \Gamma : \gamma \in \Gamma \}) \right],
\]

which proves (3.2). \( \Box \)

Now let \( \Delta_0 \) be connected in \( G_0 \). Let us say that a contour \( \gamma \) cuts \( \Delta_0 \) if \( \gamma \) contains an edge that separates some pair of vertices \( v, w \in \Delta_0 \) that are adjacent in \( G_0 \). Then, obviously, the event that \( \Delta_0 \) is uniformly colored corresponds to the event that no contour \( \gamma \in \Gamma \) cuts \( \Delta_0 \). Let \( \nu_{\Lambda} \mid \Delta_0 \) denote the measure \( \nu_{\Lambda} \) from (3.1) conditioned on this event. Let \( S_{\Delta_0}(\Gamma) \) denote the number of contours in a contour configuration \( \Gamma \) that surround \( \Delta_0 \). We obtain a lower bound for the conditional probability in (2.3) by writing

\[
\mu_{\Lambda,\infty}^{\Gamma}(J_{1,\Delta_0} \mid J_{\Delta_0}) \geq \nu_{\Lambda} \mid \Delta_0 (\{ \Gamma : S_{\Delta_0}(\Gamma) = 0 \})
\]

\[
\geq 1 - \sum_{\Gamma} \nu_{\Lambda} \mid \Delta_0 (\Gamma) S_{\Delta_0}(\Gamma) = 1 - \sum_{\gamma : \text{Int}(\gamma) \supseteq \Delta_0} \nu_{\Lambda} \mid \Delta_0 (\{ \Gamma : \gamma \in \Gamma \}).
\]

Then the probability that \( \Gamma \) contains a given contour \( \gamma \) is bounded under \( \nu_{\Lambda} \mid \Delta_0 \) in exactly the same way as it was bounded under \( \nu_{\Lambda} \) [cf. (3.3)], yielding

\[
\nu_{\Lambda} \mid \Delta_0 (\{ \Gamma : \gamma \in \Gamma \}) \leq \frac{2^{1-|\gamma|}}{1 + 2^{1-|\gamma|}}.
\]
for every $\gamma$ that surrounds $\Delta_0$. Inserting this into (3.4) yields:

**Lemma 3.2. (Peierls bound for zero temperature)** Let $\Lambda \subset V$ be a simply connected finite set such that $\partial \Lambda \subset V_0$, and let $\Delta_0 \subseteq \Lambda \cap V_0$ be connected in $G_0$. Then

$$1 - \mu_{1,\infty}^1(\mathcal{J}_1,\Delta_0 \mid \mathcal{J}_{\Delta_0}) \leq \sum_{\gamma : \text{Int}(\gamma) \supseteq \Delta_0} \frac{2^{1-|\gamma|}}{1 + 2^{1-|\gamma|}} = \sum_{L=3}^{\infty} N_{\Delta_0}(L) \frac{2^{1-L}}{1 + 2^{1-L}},$$

where $N_{\Delta_0}(L)$ denotes the number of contours of length $L$ surrounding $\Delta_0$.

Our proofs of Lemmas 2.1 and 2.5 in the zero-temperature case will be based on the estimate (3.6) and suitable bounds on the numbers $N_{\Delta_0}(L)$.

**Remarks.**

1. In the special case that $\Delta_0$ is a singleton, the event $\mathcal{J}_{\Delta_0}$ is trivially fulfilled and the conditional probability in (3.6) reduces to an unconditional probability.

2. The simpler but slightly weaker bound

$$1 - \mu_{1,\infty}^1(\mathcal{J}_1,\Delta_0 \mid \mathcal{J}_{\Delta_0}) \leq \sum_{L=3}^{\infty} N_{\Delta_0}(L) 2^{1-L}$$

is sufficient for nearly all purposes. Indeed, even for quantitative bounds the difference between (3.6) and (3.7) is very small: for instance, when $G$ is the diced lattice and $G_1$ is the hexagonal lattice, we have $L \geq 6$, so one sees immediately that the difference between (3.6) and (3.7) cannot be more than about 3%. See also the proof of Lemma 2.5 for $\beta = \infty$ in Section 3.4 below.

**3.2. Bounds on contours for zero temperature.** The main ingredient in the proof of Lemma 2.1 will be a bound on the number of simple circuits in $G_1$ of a given length surrounding a given vertex in $G_0$. We start by bounding the number of self-avoiding paths in $G_1$, or more generally in quasi-transitive graphs of bounded degree. We then use this bound to obtain a bound on self-avoiding polygons, i.e. simple circuits.

Let $H = (V, E)$ be any graph. It will be convenient to view $H$ as a directed graph, by introducing a pair of directed edges (one in each direction) corresponding to each edge of the undirected graph $H$. So let $A$ be the set of directed edges of $H$, i.e., $A$ is the set of all ordered pairs $(v, w)$ of vertices such that $(v, w) \in E$. By definition, a self-avoiding path in $G$ of length $n$ is a finite sequence of vertices $v_0, \ldots, v_n \in V$, all different from each other, such that $(v_{k-1}, v_k) \in A$ for all $k = 1, \ldots, n$. We call $(v_0, v_1)$ the starting edge and $(v_{n-1}, v_n)$ the final edge of the path. For $n \geq 1$ and $a, b \in A$, we denote by $C_n(a, b)$ the number of self-avoiding paths in $G$ of length $n$ with starting edge $a$ and final edge $b$. We then set

$$C_n(a) := \sum_{b \in A} C_n(a, b) \quad \text{and} \quad C_n^* := \sup_{a \in A} C_n(a).$$
Lemma 3.3. (Exponential bound on self-avoiding paths) Let $H = (V,E)$ be an infinite connected graph in which each vertex has degree at most $k$. Then the limit
\begin{equation}
\alpha(H) := \lim_{n \to \infty} (C_n^*)^{1/n}
\end{equation}
exists and equals $\inf_{n \geq 1} (C_{n+1}^*)^{1/n}$; it satisfies $1 \leq \alpha(H) \leq k - 1$. Furthermore, if $H$ is quasi-transitive and is anything other than a tree in which every vertex has degree $k$, then $\alpha(H) < k - 1$.

Proof. For $m, n \geq 1$ and $a,c \in A$ we have
\begin{equation}
C_{m+n-1}(a,c) \leq \sum_{b \in A} C_m(a,b) C_n(b,c)
\end{equation}
because any self-avoiding path of length $m+n-1$ can be decomposed uniquely into its first $m$ steps and its last $n$ steps, each of which is a self-avoiding path, which overlap in a single directed edge (here called $b$). This implies the submultiplicativity
\begin{equation}
C_{m+n}^* \leq C_m^* C_n^*.
\end{equation}
We see that $n \mapsto \log C_{n+1}^*$ $(n \geq 0)$ is subadditive, which implies (see, e.g., Theorem B.22) that the limit
\begin{equation}
\alpha(H) := \lim_{n \to \infty} (C_{n+1}^*)^{1/n} = \inf_{n \geq 1} (C_{n+1}^*)^{1/n}
\end{equation}
exists, with $0 \leq \alpha(H) < \infty$.

By Lemma A.2(a), there exists an infinite self-avoiding path $(v_0,v_1,v_2,\ldots)$; so taking $a = (v_0,v_1)$ we see that $C_n(a) \geq 1$ for all $n \geq 1$. Hence $\alpha(H) \geq 1$.

Since each $v \in V$ is of degree at most $k$, self-avoidance trivially implies that
\begin{equation}
C_{n+1}^* \leq (k - 1)^n,
\end{equation}
so that $\alpha(H) \leq k - 1$.

If $H$ is anything other than a $k$-regular tree, then since $H$ is connected, for each $a \in A$ there exists an integer $m$ (depending only on the equivalence class of $a$ under the automorphism group of $H$) such that $C_{m+1}^*(a) < (k - 1)^m$: it suffices to walk to a vertex of degree $< k$ and then one step more, or else walk into and around a circuit. Using the submultiplicativity (3.11) together with (3.13), it follows that $C_{n+1}^*(a) < (k - 1)^n$ for all $n \geq m$. If now $H$ is (vertex-)quasi-transitive, then it is not hard to see that it is also directed-edge-quasi-transitive (see Lemma A.1 for a proof), i.e. there are finitely many equivalence classes of directed edges, so we can choose an $m$ that works for all $a \in A$. It follows that $C_{n+1}^* < (k - 1)^n$ for some $n$ (in fact for all sufficiently large $n$), which shows that the infimum in (3.12) is strictly less than $k - 1$. □

Remark. Most of this proof can alternatively be carried out in terms of the more familiar vertex-to-vertex counts $c_n(u,v)$ for $n \geq 0$ and the corresponding quantities
\[ c_n^* = \sup_{u \in V} \sum_{v \in V} c_n(u, v). \]  (Since \( C_n^* \leq c_n^* \leq kC_n^* \), the two counts have identical asymptotic growth.) Indeed for \( m, n \geq 0 \) we have
\[
(3.14) \quad c_{m+n}(u, w) \leq \sum_{v \in V} c_m(u, v) c_n(v, w)
\]
and hence \( c_{m+n}^* \leq c_m^* c_n^* \), from which it follows that
\[
(3.15) \quad \alpha(H) = \lim_{n \to \infty} (c_n^*)^{1/n} = \inf_{n \geq 1} (c_n^*)^{1/n}
\]
exists. But it is more difficult in this framework to prove that \( \alpha(H) < k^{-1} \), since the bound \( c_n^* \leq k(k-1)^{n-1} \) has an extra factor \( k/(k-1) \) that we must somehow overcome. It is for this reason that we found it convenient to work with directed edges instead of vertices.

It follows from (3.9) that for each \( \epsilon > 0 \) there exists \( K_\epsilon < \infty \) such that
\[
(3.16) \quad C_n^* \leq K_\epsilon [\alpha(H) + \epsilon]^n \quad \text{for all } n \geq 0.
\]

Now let \( G = (V, E) \) be as in Theorem 1.1 and let \( G_0 = (V_0, E_0) \) and \( G_1 = (V_1, E_1) \) be its sublattices. Recall from Section 3.1 that \( N_{\Delta_0}(L) \) denotes the number of simple circuits of length \( L \) in \( G_1 \) surrounding a set \( \Delta_0 \subset V_0 \). Lemma 3.3 applied to \( H = G_1 \) implies the following bound on \( N_{\Delta_0}(L) \):

**Lemma 3.4. (Exponential bound on circuits surrounding a point)** We have \( \alpha(G_1) < 2 \). Moreover, for every \( \epsilon > 0 \), there exists a constant \( C_\epsilon < \infty \) such that
\[
(3.17) \quad N_{\{v\}}(L) \leq C_\epsilon [\alpha(G_1) + \epsilon]^L
\]
for all \( v \in V_0 \) and all \( L \geq 1 \).

**Proof.** Since every vertex in \( G_1 \) has degree 3 and \( G_1 \) is not a tree (indeed, each vertex in \( G_0 \) is surrounded by a circuit in \( G_1 \)), it follows from Lemma 3.3 that \( \alpha(G_1) < 2 \).

Since \( G \) is infinite, connected and locally finite, it is not hard to show [see Lemma A.2(a) in the Appendix] that for each \( v \in V_0 \) we can find an infinite self-avoiding path \( \pi = (v_0, v_1, \ldots) \) in \( G_0 \) starting at \( v_0 = v \) such that the graph distance (in \( G_0 \)) of \( v_n \) to \( v \) is \( n \). It is not hard to see that any simple circuit surrounding \( v \) must cross some edge of \( \pi \). With a bit more work, we can get a quantitative bound on how far this edge can be from the starting point of \( \pi \). Indeed, it follows from Proposition A.5 that there exists a constant \( K < \infty \), depending only on the graph \( G_0 \), such that any simple circuit of length \( L \) surrounding \( v \) must cross one of the first \( N \) edges of \( \pi \), where
\[
(3.18) \quad N := 1 + K + \frac{1}{2} \left( \frac{3}{2} - 1 \right) L = 1 + K + L/4.
\]

So let \( \gamma \) be a simple circuit of length \( L \) surrounding \( v \). Let \((v_{k-1}, v_k)\) be the first edge of \( \pi \) that is crossed by \( \gamma \), and let \( a \) be the corresponding (dual) edge in \( \gamma \). We can view \( a \) as a directed edge by agreeing that we turn \((v_{k-1}, v_k)\) anticlockwise to get...
a. Then we can specify $\gamma$ completely by specifying the first edge of $\pi$ to be crossed by $\gamma$ and by specifying the self-avoiding path formed by the first $L - 1$ edges of $\gamma$, starting with $a$. By (3.18), this yields the bound

$$N(v)(L) \leq (1 + K + L/4) C_{L-1}^*.$$

By (3.16), the claim follows: it suffices to absorb the factor $(1 + K + L/4)$ into a change of the base of the exponential term. □

3.3. Long-range dependence for zero temperature. We are now ready to prove Lemma 2.1 for zero temperature.

**Proof of Lemma 2.1 for $\beta = \infty$.** It follows from Proposition A.5 that for each $L_0 < \infty$, there exists $M < \infty$ such that each finite, $G_0$-connected set $\Delta_0 \subset V_0$ with $|\Delta_0| \geq M$ has the property that any simple circuit in $G_1$ surrounding $\Delta_0$ must be of length at least $L_0$.

Then the weak Peierls bound (3.7) and Lemma 3.4 imply that for any $\epsilon > 0$ there exists $C_\epsilon < \infty$ such that for every finite and simply connected $\Lambda \supset \Delta_0$ with $\partial \Lambda \subset V_0$, we have

$$\mu^1_{\Lambda,\infty}(J_{1,\Delta_0} | J_{\Delta_0}) \geq 1 - C_\epsilon \sum_{L=L_0}^{\infty} 2^{1-L}[\alpha(G_1) + \epsilon]^L.$$

Since $\alpha(G_1) < 2$, by choosing first $\epsilon$ small enough and then $L_0$ large enough (and $M$ appropriately), we can make the conditional probability in (3.20) as close to 1 as we wish, uniformly in $\Lambda$. □

3.4. Quantitative bound for the diced lattice. Let $p_L$ denote the number of simple circuits (i.e., self-avoiding polygons) of length $L$ in the hexagonal lattice, modulo translation. And let $q_L$ denote the number of simple circuits of length $L$ in the hexagonal lattice that surround a given vertex of the triangular lattice. We have the following bounds:

**Lemma 3.5.** (Supermultiplicativity of hexagonal-lattice polygons) The number $p_L$ of hexagonal-lattice self-avoiding polygons of length $L$, modulo translation, satisfies

$$p_{L+M-2} \geq p_L p_M.$$

**Corollary 3.6.** (Bound on hexagonal-lattice circuits surrounding a point) The number of simple circuits in the hexagonal lattice $G_1$ surrounding a given vertex in $G_0$ is bounded as

$$q_L \leq (L^2/36) (2 + \sqrt{2})^{(L-2)/2}.$$
Proof of Lemma 3.5. We use concatenation: Consider two polygons \(\gamma_1\) and \(\gamma_2\) contributing to \(p_L\) and \(p_M\), respectively. Let \((x, x + e_2)\) be the highest vertical edge of \(\gamma_1\) in its rightmost column, and let \((y, y + e_2)\) be the lowest vertical edge of \(\gamma_2\) in its leftmost column, where \(e_1 := (1, 0)\) and \(e_2 := (0, 1)\) denote the natural basisvectors of \(\mathbb{R}^2\). Uniting the polygon \(\gamma_2\) with \(\gamma_1\) shifted by \(y - x\) and erasing the edges \((y, y + e_2)\), we get a contour \(\gamma = T_{y-x}(\gamma_1) \cup \gamma_2 \setminus (y, y + e_2)\) contributing to \(p_{L+M-2}\). To complete the argument, we must show that different choices of \(\gamma_1\) and/or \(\gamma_2\) lead to a different \(\gamma\) (modulo translation), i.e., we can reconstruct \(\gamma_1\) and \(\gamma_2\) (modulo translation) from \(\gamma\). To this aim, we observe that \((y, y + e_2)\) is the only vertical edge in its column that cuts the interior of \(\gamma\). Also, if another column cuts the interior of \(\gamma\) in a single edge, then the contours \(\gamma_1'\) and \(\gamma_2'\) obtained by cutting \(\gamma\) at this edge into a left and right piece will have lengths different from \(L\) and \(M\). Thus, for fixed \(L\) and \(M\), each different (modulo translations) ordered pair \((\gamma_1, \gamma_2)\) of polygons of lengths \(L\) and \(M\) yields a different (modulo translations) polygon of length \(L + M - 2\). □

Proof of Corollary 3.6. The proof combines three ingredients. The first is the fact, conjectured in [56] and proven in [21], that the connective constant of the hexagonal lattice is exactly \(\alpha = \sqrt{2 + \sqrt{2}} \approx 1.847759\). The second ingredient is the isoperimetric inequality for the hexagonal lattice: the number of faces surrounded by a circuit of length \(L\) is at most \(L^2/36\). The third ingredient is a bound on the number \(p_L\) of \(L\)-step hexagonal-lattice self-avoiding polygons modulo translation in terms of the connective constant \(\alpha\) for self-avoiding walks on the hexagonal lattice, namely [47]

\[
p_L \leq \alpha^{L-2}.
\]

Indeed, the supermultiplicativity \(p_{L+M-2} \geq p_L p_M\) implies, by standard arguments, that \(\alpha_{SAP} = \lim_{L \to \infty} (p_L)^{1/L}\) exists and that \(p_L \leq (\alpha_{SAP})^{L-2}\). On the other hand, since \(p_L \leq c_{L-1}/(2L)\) where \(c_n\) is the number of \(n\)-step self-avoiding paths starting at a given vertex, we manifestly have \(\alpha_{SAP} \leq \alpha\). □

Remarks. 1. The supermultiplicativity \(p_{L+M-2} \geq p_L p_M\) for the hexagonal lattice is stronger than the inequality \(p_{L+M} \geq p_L p_M\) that holds for the square lattice [53, Theorem 3.2.3]. As a consequence, we are able to prove \(p_L \leq \alpha^{L-2}\) rather than just \(p_L \leq \alpha^L\).

2. For self-avoiding paths and polygons on \(\mathbb{Z}^d\) it is known [53, Corollary 3.2.5] that \(\alpha_{SAP} = \alpha\). The same presumably holds also for the hexagonal lattice and for other lattices periodically embedded in Euclidean space, but we are not aware of any proof. Since we need only an upper bound on \(\alpha_{SAP}\), we have refrained from addressing this question. Note also that \(\alpha_{SAP} < \alpha\) on hyperbolic lattices (with the possible exception of eight such lattices) [54], so the equality \(\alpha_{SAP} = \alpha\) is a somewhat delicate matter.
Proof of Lemma 2.5 for \( \beta = \infty \). We use the explicit values of \( q_L \) for \( L = 6, 8, \ldots, 140 \) obtained by Jensen’s computer-assisted enumerations \[41\] together with the bound (3.22) for even \( L \geq 142 \). From \[41\] we get

\[
\sum_{L=6}^{140} q_L 2^{-L} = \frac{22074233899340881133583692519761872405249}{2^{139}} < 0.03168 .
\]

On the other hand, we have

\[
\sum_{\text{even } L \geq 142} (L^2/36)(2 + \sqrt{2})^{(L-2)/2} 2^{-L} = \frac{(2 + \sqrt{2})^{70} (2907 + 1531 \sqrt{2})}{9 \cdot 2^{139}} < 0.01731 .
\]

Putting these together, we have

\[
\sum_{L=6}^{\infty} q_L 2^{-L} < 0.04899 .
\]

Inserting this into the weak Peierls bound (3.7) specialized to \( \Delta_0 = \{v\} \), we obtain

\[
\mu_{\Lambda, \infty}^1(\sigma_v = 1) > 1 - 2(0.04899) = 0.90202 .
\]

A slight improvement of (3.27) can be obtained by using (3.6) in place of (3.7): we have

\[
\sum_{L=6}^{140} q_L \frac{2^{-L}}{1 + 2^{1-L}} < 0.03119 .
\]

(The improvement in the tail sum \( L \geq 142 \) is of course utterly negligible.) The final result (3.27) is then improved from 0.90202 to 0.90301. \( \square \)

Remarks. 1. Jensen \[41\] conjectured, based on his enumerations for \( L \leq 140 \), that the large-\( L \) asymptotic asymptotic behavior of \( q_L \) is

\[
q_L = \frac{1}{4\pi} \big(2 + \sqrt{2}\big)^{L/2} L^{-1} \big[1 + o(1)\big] .
\]

(At \( L = 140 \) the exact value for \( q_L \) is already within 0.4% of this asymptotic form.) Using this formula in place of the bound (3.22), we find for the tail

\[
\sum_{\text{even } L \geq 142} q_L 2^{-L} \approx 4.7 \times 10^{-8} \ (\ll 0.01731) .
\]

It follows that if we could know \( q_L \) exactly for all \( L \), then our Peierls argument using (3.6) would be capable of proving a lower bound 0.93762 in (3.27). This should be compared with the actual zero-temperature value 0.957597 \( \pm 0.000004 \) obtained by Monte Carlo simulations \[47\] \[18\].

\[17\] The relevant series is called there the “first area-weighted moment” for honeycomb-lattice polygons and is contained in the file \texttt{hcsapmom1.ser}.

\[18\] See footnote 8 above.
2. When [47] was written, the exact result \( \alpha = \sqrt{2 + \sqrt{2}} \approx 1.847759 \) was not yet a rigorous theorem, so we used instead the bound \( \alpha < 1.868832 \) due to Alm and Parviainen [1]. Then, to get a sufficient final estimate, the additional factor \( \alpha^{-2} \) from the improved bound (3.23) implied by stronger supermultiplicativity (see Remark 1 after the proof of Corollary 3.6) was crucial.

4. THE POSITIVE-TEMPERATURE CASE

In this section we extend the Peierls argument to positive temperature, allowing us to complete the proof of Lemmas 2.1 and 2.5. We also prove the technical Lemmas 2.2 and 2.4.

4.1. Contour model for positive temperature. As before, we consider a graph \( G = (V,E) \) satisfying the conditions of Theorem 1.1 and take a finite and simply connected set \( \Lambda \subset V \) such that \( \partial \Lambda \subset V_0 \). Our aim is to derive bounds on the probabilities of certain events under the finite-volume Gibbs measures \( \mu_{\Lambda,\beta} \) which correspond to uniform color-1 boundary conditions on \( \partial \Lambda \).

We recall from Section 2.1 that every color configuration \( \sigma \) on \( \Lambda \) (2.1)/(2.2) a collection \( E_1(\sigma) \) of edges in the sublattice \( G_1 \) that separate differently colored vertices in \( G_0 \) (or equivalently faces in \( G_1 \)). Since \( \partial \Lambda \subset V_0 \) and \( \partial \Lambda \) is uniformly colored (in color 1), each edge of \( E_1(\sigma) \) has both its endpoints in \( V_1 \cap \Lambda \). In general, we define contours to be connected components of \( E_1(\sigma) \). [More precisely, we define contours to be the connected components of the graph \((V_1 \cap \Lambda, E_1(\sigma))\) other than isolated vertices.] If \( \sigma \) is a ground state, then at each \( v \in V_1 \cap \Lambda \), either zero or two edges of \( E_1(\sigma) \) are incident, hence the connected components of \( E_1(\sigma) \) are simple circuits in \( G_1 \). But for general color configurations \( \sigma \), the connected components of \( E_1(\sigma) \) may be more complicated. In particular, it is possible that three edges of \( E_1(\sigma) \) are incident to a vertex \( v \in V_1 \cap \Lambda \). Recall that a connected graph is called bridgeless (or 2-edge-connected) if it contains no bridges (i.e., single edges the removal of which disconnects the graph). We observe that for any color configuration \( \sigma \), the connected components of \( E_1(\sigma) \) must be bridgeless, since otherwise there would be a uniformly colored region of \( G_0 \) that bounds such a bridge on both sides, contradicting the definition of \( E_1(\sigma) \).

In view of this, in the positive-temperature model let us define a contour to be a finite connected bridgeless subgraph \( \gamma \) of \( G_1 \) containing at least one edge. Note that each vertex of such a contour has degree 2 or 3. It is easy to see that the number of vertices of degree 3 must be even (just notice that twice the number of vertices of degree two plus three times the number of vertices of degree three equals twice the number of edges). We let \( |\gamma| \) denote the number of edges of \( \gamma \), to which we will refer as the length of \( \gamma \). We let \( t(\gamma) \) be the number such that \( \gamma \) has \( 2t(\gamma) \) vertices of degree 3. Then \( \gamma \) has \( |\gamma| - 3t(\gamma) \) vertices of degree 2. Moreover, \( \gamma \) divides \( V_0 \) into \( 2 + t(\gamma) \) connected components, of which one is infinite and the others are finite and simply connected. We say that a contour \( \gamma \) surrounds a set \( \Delta_0 \subset V_0 \), denoted as \( \gamma \circ \Delta_0 \), if \( \Delta_0 \) is contained in one of the finite components. We call
the infinite component the exterior $\text{Ext}(\gamma)$ of $\gamma$, and we refer to the union of all the finite components as the interior $\text{Int}(\gamma)$ of $\gamma$. [Please note that saying that $\gamma$ surrounds a set $\Delta_0$ is stronger than saying that $\Delta_0 \subseteq \text{Int}(\gamma)$, since “surrounding” is defined as $\Delta_0$ being entirely contained in one of the finite components.] Given that the exterior of $\gamma$ is colored in one particular color, we let $\chi(\gamma)$ denote the number of possible three-colorings of the connected components of $\text{Int}(\gamma)$ in such a way that along each edge of $\gamma$, two different colors meet.\footnote{Otherwise put, $\chi(\gamma)$ is $\frac{1}{3}$ times the number of proper 3-colorings of the dual graph $\gamma^*$.} Please note that it is possible to have $\chi(\gamma) = 0$: see Figure\textsuperscript{3}. Obviously, such contours are “not allowed”, and we shall soon see that their probability is zero.\footnote{We could, if we wanted, redefine the term “contour” to include only those having $\chi(\gamma) > 0$. But there is little to be gained from complicating the definition in this way, since our counting of contours (Lemma\textsuperscript{4.2} below) is too crude to distinguish between those having $\chi(\gamma) > 0$ or $\chi(\gamma) = 0$.}

Finally, let us observe that $\chi(\gamma) \leq 2^{t(\gamma)+1}$.

We now claim that if $\sigma$ is distributed according to $\mu_{\Lambda,\beta}^1$, and $\Gamma(\sigma)$ is the collection of connected components of $(V_1 \cap \Lambda, E_1(\sigma))$ other than isolated vertices, then $\Gamma(\sigma)$ is distributed according to the law

$$\nu_{\Lambda,\beta}(\Gamma) = \frac{1}{Z_{\Lambda,\beta}} \prod_{\gamma \in \Gamma} \chi(\gamma) \beta^{|\gamma|} q_\beta^{t(\gamma)}, \tag{4.1}$$

Figure 3. A contour $\gamma$ with $\chi(\gamma) = 0.$
where $Z_{\Lambda, \beta}$ is a normalizing constant and

\begin{align}
  p_\beta &:= \frac{1 + e^{-\beta} + e^{-2\beta}}{2 + e^{-3\beta}} \\
  q_\beta &:= \frac{9e^{-2\beta}(2 + e^{-3\beta})}{(1 + e^{-\beta} + e^{-2\beta})^3}
\end{align}

(4.2a) (4.2b)

To see this, note that there are $\prod_{\gamma \in \Gamma} \chi(\gamma)$ ways of coloring the sites in $V_0 \cap \Lambda$ in a way that is consistent with $\Gamma$. Given a coloring of $V_0 \cap \Lambda$, summing the probabilities of all possible colorings of $V_1 \cap \Lambda$ yields for each site in $V_1 \cap \Lambda$ a factor

\begin{equation}
1 + 1 + e^{-3\beta}, \quad 1 + e^{-\beta} + e^{-2\beta} \quad \text{or} \quad e^{-\beta} + e^{-\beta} + e^{-\beta}
\end{equation}

(4.3)

depending on whether the site has neighbors with one, two or three different colors, respectively. These cases correspond, respectively, to sites not on a contour, sites of degree 2 on a contour, and sites of degree 3 on a contour. Absorbing the factor $1 + 1 + e^{-3\beta}$ into the normalization constant $Z_{\Lambda, \beta}$, we get a factor $\left(1 + e^{-\beta} + e^{-2\beta}\right) / (2 + e^{-3\beta})$ for each of the $|\gamma| - 3t(\gamma)$ sites of degree 2, and a factor $3e^{-\beta} / (2 + e^{-3\beta})$ for each of the $2t(\gamma)$ sites of degree 3. Putting this all together, we arrive at (4.1)/(4.2).

In the limit $\beta \to \infty$, we have $p_\beta \to \frac{1}{2}$ and $q_\beta \to 0$; in particular, the only contours that get nonzero weight in this limit are simple circuits, for which $\chi(\gamma) = 2$. Then the contour law (4.1) reduces to (3.1), as expected.

More generally, it can be easily verified that $p_\beta$ decreases monotonically from 1 to $\frac{1}{2}$ as $\beta$ runs from 0 to $\infty$, and behaves for large $\beta$ as $\frac{1}{2} + O(e^{-\beta})$; and that $q_\beta$ decreases monotonically from 1 to 0 as $\beta$ runs from 0 to $\infty$, and behaves for large $\beta$ as $O(e^{-2\beta})$.

By the same arguments as in (3.4)–(3.5), and using $\chi(\gamma) \leq 2^{t(\gamma)+1}$, we find:

Lemma 4.1. (Peierls bound for positive temperature) Let $\Lambda \subset V$ be a simply connected finite set such that $\partial \Lambda \subset V_0$, and let $\Delta_0 \subset \Lambda \cap V_0$ be connected in $G_0$. Then

\begin{equation}
1 - \mu^{1}_{\Lambda, \beta}(J_{1, \Delta_0} \mid J_{\Delta_0}) \leq \sum_{t=0}^{\infty} \sum_{L=3}^{\infty} N_{\Delta_0}(L, T) \frac{2^{T+1} p_{\beta}^{L} q_{\beta}^{T}}{1 + 2^{T+1} p_{\beta}^{L} q_{\beta}^{T}},
\end{equation}

(4.4)

where $N_{\Delta_0}(L, T)$ denotes the number of contours $\gamma$ surrounding $\Delta_0$ satisfying $|\gamma| = L$ and $t(\gamma) = T$.

4.2. Bounds on contours for positive temperature. In this section, we prove Lemmas 2.1 and 2.5. We first need to generalize Lemma 3.4 to contours that are not simple circuits. Recall that $N_{\Delta_0}(L, T)$ denotes the number of contours $\gamma$ surrounding $\Delta_0$ satisfying $|\gamma| = L$ and $t(\gamma) = T$. Recall also from Lemma 3.3 that $\alpha(H)$ denotes the connective constant of a graph $H$ as defined in (3.9), and from Lemma B.3 that $\alpha(G_1) < 2$. 

...
 Lemma 4.2. (Bound on number of contours) For every $\epsilon > 0$ there exists a constant $C'_\epsilon < \infty$ such that

$$N(v)(L,T) \leq (L^T/T!)^2 (C'_\epsilon)^{T+1} \alpha(G_1) + \epsilon^L$$

for all $v \in V_0$ and all $L, T \geq 0$.

Proof. Let $\gamma$ be a contour surrounding $\{v\}$ such that $|\gamma| = L$ and $t(\gamma) = T$. We need a suitable way to encode $\gamma$. We begin, as in the proof of Lemma 3.4, by letting $\pi = (v_0, v_1, \ldots)$ be an infinite self-avoiding path in $G_0$ starting at $v_0 = v$ such that the graph distance (in $G_0$) of $v_n$ to $v$ is $n$. According to Proposition A.5 and formula (3.18), the contour $\gamma$ intersects an edge $(v_{b-1}, v_b)$ of $\pi$ with $b \leq N := 1 + K + L/4$, where $K$ is a constant depending only on the graph $G_0$. Thus, we can find some directed simple circuit $\gamma^* = (u_1, \ldots, u_{n_1}, u_1)$ contained in $\gamma$, such that $(u_1, u_2)$ crosses the edge $(v_{b-1}, v_b)$ in the anticlockwise direction (see Figure 4).

Let us write $\gamma^1 = (u_1, \ldots, u_{n_1})$, which is a self-avoiding path. If $T = 0$, then $\gamma = \gamma^*$ and our encoding is complete. Otherwise, let $s_1 := \min\{i \geq 1: u_i$ is of degree 3
in $\gamma$. Then we can find a self-avoiding path
\begin{equation}
\gamma^2 = (u_{s_1}, u_{n_1+1}, \ldots, u_{n_2}, u'_{s_1})
\end{equation}
in $\gamma$ such that only the starting and ending points $u_{s_1}$ and $u'_{s_1}$ are in $\gamma^1$. If $T = 1$, then $\gamma = \gamma^* \cup \gamma^2$ and we are done. Otherwise, let $s_2 := \min\{i > s_1: i \neq s'_1 \text{ and } u_i \text{ is of degree } 3 \text{ in } \gamma\}$. Then we can find another self-avoiding path
\begin{equation}
\gamma^3 = (u_{s_2}, u_{n_2+1}, \ldots, u_{n_3}, u'_{s_2})
\end{equation}
in $\gamma$ such that only the starting and ending points $u_{s_2}$ and $u'_{s_2}$ are in $\gamma^1 \cup \gamma^2$. Continuing in this way, we see that we can code all the information needed to construct $\gamma$ by specifying numbers
\begin{equation}
b \leq N, \quad 2 = n_0 < n_1 < \cdots < n_{T+1} = L - T \quad \text{and} \quad 0 < s_1 < \cdots < s_T < L - T
\end{equation}
and self-avoiding paths $\gamma^1, \ldots, \gamma^{T+1}$ of lengths $n_1 - n_0 + 1, n_2 - n_1 + 1, \ldots, n_{T+1} - n_T + 1$ whose starting edges are uniquely determined by the information previously coded. By Lemma 3.3 and its consequence (3.16), for each $\epsilon > 0$ there exists $K_\epsilon < \infty$ such that the number of self-avoiding paths of length $n$ with a specified starting edge is bounded from above by $K_\epsilon [\alpha(G_1) + \epsilon]^n$.

Therefore, there are at most
\begin{equation}
\prod_{i=1}^{T+1} K_\epsilon [\alpha(G_1) + \epsilon]^{n_i - n_{i-1} + 1} = (K_\epsilon)^{T+1} [\alpha(G_1) + \epsilon]^{L-1}
\end{equation}
different contours $\gamma$ associated with given data $b, n_1, \ldots, n_T, s_1, \ldots, s_T$. Since there are $\binom{L - T - 3}{T}$ and $\binom{L - T - 1}{T}$ ways of choosing numbers $2 < n_1 < \cdots < n_T < L - T$ and $0 < s_1 < \cdots < s_T < L - T$, respectively, and since $b \leq N = 1 + K + L/4$, summing over all ways to choose the numbers $b, n_1, \ldots, n_T, s_1, \ldots, s_T$ shows that the total number of contours $\gamma$ surrounding $v$ with given $|\gamma| = L$ and $t(\gamma) = T$ is bounded by
\begin{equation}
(1 + K + L/4) \binom{L - T - 3}{T} \binom{L - T - 1}{T} (K_\epsilon)^{T+1} [\alpha(G_1) + \epsilon]^{L-1}
\leq (C'_\epsilon)^{T+1} \binom{L - T}{T}^2 [\alpha(G_1) + 2\epsilon]^{L-1} \leq (C'_\epsilon)^{T+1} (L^T / T!)^2 [\alpha(G_1) + 2\epsilon]^L,
\end{equation}
where the factor $1 + K + L/4$ was absorbed into a change of base of the exponential term followed by the change of constant into $C'_\epsilon$. \qed
4.3. Long-range dependence for positive temperature.

Proof of Lemmas 2.4 and 2.5 in the positive-temperature case. In the zero-temperature case, both lemmas have already been proven in Sections 3.3 and 3.4, respectively, by showing that for some sufficiently large \( \Delta_0 \) (respectively \( \Delta_0 = \{v\} \)) the right-hand side of (3.6) can be made sufficiently small. To generalize the two lemmas to small positive temperatures, it therefore suffices to show that the right-hand side of (4.4) converges to the right-hand side of (3.6) as \( \beta \to \infty \) [for Lemma 2.5 we should also show that the error is \( O(e^{-\beta}) \)]. In view of this, Lemmas 2.4 and 2.5 are consequences of the following lemma. □

Lemma 4.3. (Large-\( \beta \) behavior of the Peierls bound) There exist \( \beta_0, C < \infty \) such that

\[
0 \leq \sum_{T=0}^\infty \sum_{L=3}^\infty N_{\Delta_0}(L, T) \frac{2^{T+1} p_\beta^L q_\beta^T}{1 + 2^{T+1} p_\beta^L q_\beta^T} - \sum_{L=3}^\infty N_{\Delta_0}(L) \frac{2^{1-L}}{1 + 2^{1-L}} \leq Ce^{-\beta} \tag{4.11}
\]

and

\[
\sum_{T=1}^\infty \sum_{L=3}^\infty N_{\Delta_0}(L, T) \frac{2^{T+1} p_\beta^L q_\beta^T}{1 + 2^{T+1} p_\beta^L q_\beta^T} \leq \sum_{T=1}^\infty \sum_{L=3}^\infty N_{\Delta_0}(L, T) 2^{T+1} p_\beta^L q_\beta^T \leq Ce^{-2\beta} \tag{4.12}
\]

uniformly for \( \beta \in [\beta_0, \infty] \) and for nonempty finite \( G_0 \)-connected sets \( \Delta_0 \subset V_0 \).

Proof. The lower bound in (4.11) is a trivial consequence of \( p_\beta \geq \frac{1}{2} \) and \( q_\beta \geq 0 \). To prove the upper bounds, we split the double sum in (4.11) into its contributions \( T = 0 \) and \( T \geq 1 \) and bound them separately, using \( p_\beta = \frac{1}{2} + O(e^{-\beta}) \) and \( q_\beta = O(e^{-2\beta}) \).

**T = 0.** By Lemma 3.4, there exist \( C < \infty \) and \( \alpha < 2 \) such that \( N_{\Delta_0}(L) \leq C \alpha^L \). The term \( T = 0 \) can therefore be bounded (using \( p_\beta \geq \frac{1}{2} \)) as

\[
\sum_{L=3}^\infty N_{\Delta_0}(L) \frac{2 p_\beta^L}{1 + 2 p_\beta^L} \leq \sum_{L=3}^\infty N_{\Delta_0}(L) \frac{2^{1-L}}{1 + 2^{1-L}} + 2C \sum_{L=3}^\infty \alpha^L p_\beta^L - \left(\frac{1}{2}\right)^L \tag{4.13}
\]

Choosing \( \beta_0 \) large enough so that \( \alpha p_{\beta_0} < 1 \), it is easy to see, using \( p_\beta = \frac{1}{2} + O(e^{-\beta}) \), that the last term in (4.13) is \( O(e^{-\beta}) \).

**T \geq 1.** By Lemmas 3.4 and 4.2, there exist \( C < \infty \) and \( \alpha < 2 \) such that \( N_{\Delta_0}(L, T) \leq (LT/T)^L C^{T+1} \alpha^L \). Therefore the terms \( T \geq 1 \) in (4.11) can be bounded
as
\[
\sum_{T=1}^{\infty} \sum_{L=3}^{\infty} N_{\Delta_0}(L, T) \frac{2^{T+1} \beta^L \eta^T}{1 + 2^{T+1} \beta^L \eta^T} \leq \sum_{T=1}^{\infty} \sum_{L=3}^{\infty} N_{\Delta_0}(L, T) 2^{T+1} \beta^L \eta^T
\]
\[
\leq 2C \sum_{L=3}^{\infty} (\alpha \beta)^L \sum_{T=1}^{\infty} \frac{(2C \beta)^T}{(T!)^2} L^{2T}
\]
\[
\leq 2C \sum_{L=3}^{\infty} (\alpha \beta)^L \sum_{T=1}^{\infty} \frac{(8C \beta)^T}{(2T)!} L^{2T}
\]
\[
= 16C^2 \beta^\infty \sum_{L=3}^{\infty} L^2 (\alpha \beta)^L \sum_{T=0}^{\infty} \frac{(8C \beta)^T}{(2T + 2)!} L^{2T}
\]
(4.14)

where we used \(\frac{(2T)!}{(T!)^2} = \frac{(2T)}{T} \leq 2^{2T}\). Choosing \(\beta_0\) large enough so that one has \(\alpha \beta_0^e \sqrt{8C \beta} < 1\), we see that (4.14) is \(O(q^\beta) = O(e^{-2\beta})\), which proves (4.12) and completes the proof of (4.11). □

The bound (4.12) from Lemma 4.3 has a useful corollary. Let us say that a contour \(\gamma\) is simple if it is a simple circuit, i.e. \(t(\gamma) = 0\). For any contour configuration \(\Gamma\) and any \(v \in V_0\), let \(S_v(\Gamma)\) denote the number of non-simple contours in \(\Gamma\) that surround \{v\}. We then have the following bound showing that non-simple contours are rare at low temperature:

**Corollary 4.4. (Rarity of non-simple contours)** Let \(\nu_{\Lambda, \beta}\) be the contour measure from (4.1). Then there exist \(\beta_0, C < \infty\) such that
\[
\sum_{\Gamma} \nu_{\Lambda, \beta}(\Gamma) S_v(\Gamma) \leq Ce^{-2\beta}
\]
(4.15)
uniformly for \(\beta \in [\beta_0, \infty]\), for finite simply connected \(\Lambda \subset V\) such that \(\partial \Lambda \subset V_0\), and for \(v \in \Lambda \cap V_0\).

**Proof.** This is an immediate consequence of (4.12) together with the positive-temperature analogue of Lemma 3.1. □

4.4. **Proof of the technical lemmas.** In this section we prove Lemmas 2.2 and 2.4.

**Proof of Lemma 2.2.** It is easy to show that for each \(\beta_0 < \infty\) there exists an \(\epsilon > 0\) such that \(\mu_{\Lambda, \beta}(\mathcal{F}_{\Delta_0}) \geq \epsilon\), uniformly for all \(0 \leq \beta \leq \beta_0\) and all finite and simply connected \(\Lambda \supset \Delta_0\) such that \(\partial \Lambda \subset V_0\). Indeed, this follows from a “finite energy”
argument: given any configuration \( \sigma \in \{1, 2, 3\}^\Lambda \), we can recolor the sites in \( \Delta_0 \) in any color of our choice at an energetic cost of at worst \( e^{-\beta|\Delta_0|} \) and an entropic cost of at worst \( 3^{-|\Delta_0|} \). Note that here \( \Delta_0 \) is fixed and finite, so the precise dependence of the costs on \( \Delta_0 \) is irrelevant. The only difficulty is that the bound one obtains in this way is not uniform in \( \beta \) as \( \beta \to \infty \). Therefore, to complete the proof, it suffices to show that there exists some \( \beta_0 < \infty \) such that \( \mu^{1, \beta}_A(J_{\Delta_0}) \) can be estimated from below uniformly in \( \beta_0 \leq \beta \leq \infty \) and \( \Lambda \).

In order to prove this, let \( \Delta_0 \subset V_0 \) be finite and \( G_0 \)-connected, and let \( \overline{\Delta_0} \) denote the union of \( \Delta_0 \) with its boundary in \( G_0 \), i.e., \( \overline{\Delta_0} := \Delta_0 \cup \{ v \in V_0 : \exists u \in \Delta_0 \text{ s.t. } \{u, v\} \in E_0 \} \). By Corollary 4.4, the probability that a non-simple contour intersects \( \overline{\Delta_0} \) tends to zero as \( \beta \to \infty \), uniformly in \( \Lambda \). Thus, we may choose \( \beta_0 < \infty \) such that

\[
\nu_{A, \beta}\left( \{ \Gamma : \# \gamma \in \Gamma \text{ s.t. } t(\gamma) \geq 1, \gamma \text{ intersects } \overline{\Delta_0} \} \right) \geq 1/2,
\]

uniformly in \( \Lambda \) and \( \beta_0 \leq \beta \leq \infty \). If all contours intersecting \( \overline{\Delta_0} \) are simple, then we claim that we can change our contour configuration at a finite energetic cost uniformly in \( \beta_0 \leq \beta \leq \infty \), so that no contour intersects \( \Delta_0 \). To describe the algorithm of changing a contour configuration \( \Gamma \) into a configuration \( \Gamma' \) with no contour intersecting \( \Delta_0 \), we first observe that relying on the fact that all contours intersecting \( \overline{\Delta_0} \) are simple, we can color the vertices in \( \Lambda \cap V_0 \) in three colors in such a way that boundaries between different colors correspond to contours and only two different colors occur in \( \overline{\Delta_0} \). (Note that this part of the argument uses that the contours intersecting \( \overline{\Delta_0} \) are simple everywhere and not just that there are no triple points inside \( \overline{\Delta_0} \).) Now we change our coloring by painting \( \Delta_0 \) uniformly in one of these two colors, defining thus the new contour configuration \( \Gamma' \) (see Figure 5).

Since in the construction of \( \Gamma' \) a two-color configuration in \( \overline{\Delta_0} \) was changed using the same two colors no new triple points are introduced. This, together with (4.1) and a standard finite-energy argument proves our claim.

For completeness, we write down this finite-energy argument in detail. Let \( \Gamma \) and \( \Gamma' = \psi(\Gamma) \) denote the old and new contour configuration obtained by the procedure described above. We need to estimate the relative probability of \( \Gamma' \) with respect to \( \Gamma \) and the number of different configurations \( \Gamma \) that can be mapped onto the same \( \Gamma' \), \( |\Psi^{-1}(\Gamma')| \). Let \( |\Delta_0| \) be the number of sites in \( \Delta_0 \), let \( M_{\Delta_0} \) be the number of edges in \( E_1 \) that separate sites in \( \Delta_0 \) from each other and let \( M_{\partial \Delta_0} \) be the number of edges in \( E_1 \) that separate sites in \( \Delta_0 \) from sites in \( \Delta_0 \setminus \Delta_0 \). Further, let

\[
(4.17) \quad \chi(\Gamma) := \prod_{\gamma \in \Gamma} \chi(\gamma), \quad |\Gamma| := \sum_{\gamma \in \Gamma} |\gamma| \quad \text{and} \quad t(\Gamma) := \sum_{\gamma \in \Gamma} t(\gamma).
\]

Since all contours we remove or alter are simple contours with \( \chi(\gamma) = 2 \) and we remove or alter no more than \( M_{\Delta_0} \) contours from our configuration and add no more than \( M_{\partial \Delta_0} \) edges, we have

\[
(4.18) \quad \chi(\Gamma') \geq 2^{-M_{\Delta_0}} \chi(\Gamma) \quad \text{and} \quad |\Gamma'| \leq |\Gamma| + M_{\partial \Delta_0}.
\]
Figure 5. Simple contours intersecting the square $\Delta_0$ colored with two colors: 1 (white) and 2 (gray). After flipping all sites in $\Delta_0$ to the color 1, no contour is intersecting $\Delta_0$ and no triple point was created. The same would be true when flipping all sites in $\Delta_0$ to the color 2.

while $t(\Gamma') = t(\Gamma)$, which by (4.1) implies that

$$\nu_{\Lambda,\beta}(\Gamma') \geq 2^{-M_{\Delta_0}} p_{\beta}^{M_{\Delta_0}} \nu_{\Lambda,\beta}(\Gamma).$$

Moreover, since there are $2^{|\Delta_0|}$ ways of coloring the vertices in $\Delta_0$ using only two colors, we see that there are at most $2^{|\Delta_0|}$ different contour configurations $\Gamma$ in $\Psi^{-1}(\Gamma')$. Recall that $\mathcal{J}_{\Delta_0} = \{ \Gamma: \text{ no contour in } \Gamma \text{ intersects } \Delta_0 \}$ corresponds to the event that $\Delta_0$ is uniformly colored in one color. Let $S_{\overline{\Delta_0}}$ be the event that all contours intersecting $\overline{\Delta_0}$ are simple. Then

$$\nu_{\Lambda,\beta}(\mathcal{J}_{\Delta_0}) = \sum_{\Gamma' \in \mathcal{J}_{\Delta_0}} \nu_{\Lambda,\beta}(\Gamma') \geq 2^{-|\Delta_0|} \sum_{\Gamma \in S_{\overline{\Delta_0}}} \nu_{\Lambda,\beta}(\Psi(\Gamma))$$

$$\geq 2^{-|\Delta_0| - M_{\Delta_0}} p_{\beta}^{M_{\Delta_0}} \sum_{\Gamma \in S_{\overline{\Delta_0}}} \nu_{\Lambda,\beta}(\Gamma) \geq 2^{1-|\Delta_0| - M_{\Delta_0}} p_{\beta}^{M_{\Delta_0}},$$

where we have used (4.16) in the last step. □
Proof of Lemma 2.4. Consider any $v \in V_1 \cap \Lambda$ and let $w_1, w_2, w_3 \in V_0$ be its neighbors in $G$. Then the DLR equations for the volume $\{v\}$ imply that

$$\mu_{1,\beta}(\exists i \text{ with } \sigma_v = \sigma_{w_i}, \sigma_{w_1}, \sigma_{w_2}, \sigma_{w_3})$$

(4.21)

$$= \begin{cases} e^{-3\beta} & \text{if } \sigma_{w_1} = \sigma_{w_2} = \sigma_{w_3} \\ \frac{e^{-2\beta}}{2 + e^{-3\beta}} & \text{if } |\{\sigma_{w_1}, \sigma_{w_2}, \sigma_{w_3}\}| = 2 \\ 1 + e^{-\beta} + e^{-2\beta} & \text{if } |\{\sigma_{w_1}, \sigma_{w_2}, \sigma_{w_3}\}| = 3 \end{cases}$$

Let $B := \{\sigma : |\{\sigma_{w_1}, \sigma_{w_2}, \sigma_{w_3}\}| = 3\}$ be the (“bad”) event that $w_1, w_2, w_3$ are colored in three different colors. It follows from Corollary 4.4 that

$$\mu_{1,\beta}(B) \leq C e^{-2\beta}$$

(4.22)

uniformly for $\beta \in [\beta_0, \infty]$ and for finite and simply connected $\Lambda \ni v$ such that $\partial \Lambda \subset V_0$; and by increasing $C$ we can make this hold uniformly for $\beta \in [0, \infty]$. It then follows from (4.21) and (4.22) that

$$\mu_{1,\beta}(\exists i \text{ with } \sigma_v = \sigma_{w_i}) \leq C' e^{-\beta}$$

(4.23)

uniformly for $\beta \in [0, \infty]$ and for finite and simply connected $\Lambda \ni v$ such that $\partial \Lambda \subset V_0$. □

5. Positive magnetization

In this section we prove Lemma 2.3: by using this lemma we can improve the statement that sufficiently large blocks are more likely to be uniformly colored in the color 1 than in any other color, to the “positive magnetization” statements in Theorem 1.1(a,b), which say that single vertices in the sublattices $V_0$ and $V_1$ are colored with the color 1 with a probability that is strictly larger (resp. strictly smaller) than $1/3$.

We fix an arbitrary $\beta_0 > 0$ throughout this section; our estimates will be uniform in $\beta \in [\beta_0, \infty]$. We will later also fix a finite $G_0$-connected set $\Delta_0 \subset V_0$. As in all our proofs, we work with the finite-volume Gibbs measures $\mu_{1,\beta}$, where $\Lambda \subset V$ is finite and simply connected in $G$ and satisfies $\partial \Lambda \subset V_0$. We aim to derive bounds that are uniform in such $\Lambda$ with $\Lambda \ni \Delta_0$.

Unlike what was done in the preceding subsections, we will not make use of the contour description of $\mu_{1,\beta}$, nor will we integrate out one sublattice. Rather, we will work directly with the Potts antiferromagnet on our original quadrangulation $G = (V, E)$.

Note first that the measures $\mu_{1,\beta}$ are invariant under global interchange of the colors 2 and 3. In particular, we have $\mu_{1,\beta}(\sigma_v = 2) = \mu_{1,\beta}(\sigma_v = 3)$ for all $v \in \Lambda$. Thus, to show that $\mu_{1,\beta}(\sigma_v = 1) > 1/3$ (resp. $< 1/3$), we may equivalently show
that $\mu^1_{\Lambda,\beta}(\sigma_v = 1) - \mu^1_{\Lambda,\beta}(\sigma_v = 2) > 0$ (resp. $< 0$). Because of the antiferromagnetic nature of our model, it is in fact already nontrivial to show that these quantities are nonnegative (resp. nonpositive) for $v \in V_0$ (resp. $v \in V_1$). This problem has been solved, however, in [24, Appendix A], where a first Griffiths inequality for antiferromagnetic Potts models on bipartite graphs is proven using ideas based on the cluster algorithm from [70, 71].

We will elaborate on these ideas. The main step will be to give a random-cluster representation for the law of the 1’s and 2’s conditional on the 3’s. In this representation, we will see that for $v_0 \in V_0$, the difference between the probabilities that $v_0$ is colored 1 or colored 2 equals the probability that $v_0$ percolates, i.e., that $v_0$ is in the same random cluster of 1’s and 2’s as the boundary $\partial \Lambda$. Moreover, a similar statement holds for the probability that $\Delta_0$ is uniformly colored in the color 1 minus the probability that it is uniformly colored in the color 2. Thus, by showing that both of these quantities are related to percolation of the 1’s and 2’s, we can prove that if one of them is strictly positive, then so must be the other. Note that conditioning on the positions of the 3’s would not in general be a very useful thing to do when trying to prove statements about our model, since we have no a priori knowledge of the distribution of the 3’s. Nevertheless, as we see here, it can be used to show that a certain statement that has already been proved is equivalent to another statement for which we have no direct control.

So let $G = (V, E)$ be our original quadrangulation, and let $E_\Lambda$ be the set of edges in $E$ that have at least one endvertex in $\Lambda$. We define the measure $\rho^1_{\Lambda,\beta}$ on $\{1, 2, 3\}^\Lambda \times \{0, 1\}^{E_\Lambda}$ so that the marginal distribution of $\rho^1_{\Lambda,\beta}(\sigma, \eta)$ on $\sigma$ is the Gibbs measure $\mu^1_{\Lambda,\beta}$ and so that, conditional on $\sigma$, independently for each $e \in E_\Lambda$, one has $\eta_e = 1$ with probability $p := 1 - e^{-\beta}$ if $\sigma_u, \sigma_v \in \{1, 2\}$ and $\sigma_u \neq \sigma_v$, and $\eta_e = 0$ in all other cases. That is,

$$
\rho^1_{\Lambda,\beta}(\sigma, \eta) := \frac{1}{Z^1_{\Lambda,\beta}} \exp \left[ - \beta H_\Lambda(\sigma | \tau) \right]
\times \prod_{\{u,v\} \in E_\Lambda} \left( 1_{A_{u,v}} [p 1_{\{\eta_{(u,v)} = 1\}} + (1 - p) 1_{\{\eta_{(u,v)} = 0\}}] + 1_{A^c_{u,v}} 1_{\{\eta_{(u,v)} = 0\}} \right)
$$

where $A_{u,v}$ is the event

$$
A_{u,v} := \{ \sigma_u, \sigma_v \in \{1, 2\} \text{ and } \sigma_u \neq \sigma_v \},
$$

where $A^c_{u,v}$ is its complement, $\tau$ is any spin configuration that assumes the value 1 on $\partial \Lambda$, $H_\Lambda(\sigma | \tau)$ is defined in [1.4], and $Z^1_{\Lambda,\beta}$ is the same normalizing constant as in (1.6).

Now let

$$
\Lambda^{12} := \{ v \in \Lambda \cup \partial \Lambda : \sigma_v \in \{1, 2\} \}
$$

$$
\Lambda^3 := \{ v \in \Lambda : \sigma_v = 3 \}
$$
be the sets of vertices in $\Lambda \cup \partial \Lambda$ where $\sigma$ assumes the values 1 or 2 (resp. 3), and set

\[(5.4) \quad E^{12} := \{ e \in E_\Lambda : \eta_e = 1 \} . \]

Conditionally on $\Lambda^3$, the spins $(\sigma_e)_{e \in A^{12}}$ are distributed as an antiferromagnetic Ising model, with 1 boundary conditions, on the diluted lattice $\Lambda^{12}$. Since $\Lambda^{12}$ is bipartite and the boundary conditions lie entirely on the sublattice $V_0$, we may flip the spins on the other sublattice (i.e., on $\Lambda^{12} \cap V_1$) to obtain a ferromagnetic Ising model $(\sigma'_e)_{e \in A^{12}}$ on $\Lambda^{12}$. After this flipping, the conditional joint law of $(\sigma'_e)_{e \in A^{12}}$ and $\eta$ given $\Lambda^3$ is just the standard coupling of this ferromagnetic Ising model and its corresponding random-cluster model on $\Lambda^{12}$ (see [23] and [33] Section 1.4]). (Notice that for all edges $\{u, v\}$ such that $\{u, v\} \cap \Lambda^3 \neq \emptyset$, we have $\eta_{\{u, v\}} = 0$.) Returning to the original (unflipped) spins $(\sigma_e)_{e \in A^{12}}$, we see from [33] Theorem 1.13 that, conditional on $\Lambda^3$ and $\eta$, the connected components of the graph $G^{12} = (\Lambda^{12}, E^{12})$ are independently given proper 2-colorings (with colors 1 and 2) as follows: for any component not connected to the boundary $\partial \Lambda$, each of the two proper 2-colorings arises with probability 1/2; and any component connected to the boundary is given the unique proper 2-coloring that is compatible with the boundary conditions (namely, color 1 on $V_0$ and color 2 on $V_1$). In particular, for points $v_0 \in V_0 \cap \Lambda$ one has

\[(5.5) \quad \rho^1_{\Lambda, \beta}(\sigma_{v_0} = 1 \mid \Lambda^3, \eta) = \begin{cases} 1 & \text{if } v_0 \leftrightarrow_\eta \partial \Lambda \\ \frac{1}{2} & \text{if } v_0 \in \Lambda^{12} \text{ and } v_0 \not\leftrightarrow_\eta \partial \Lambda \\ 0 & \text{if } v_0 \in \Lambda^3 \end{cases} \]

\[(5.6) \quad \rho^1_{\Lambda, \beta}(\sigma_{v_0} = 2 \mid \Lambda^3, \eta) = \begin{cases} 0 & \text{if } v_0 \leftrightarrow_\eta \partial \Lambda \\ \frac{1}{2} & \text{if } v_0 \in \Lambda^{12} \text{ and } v_0 \not\leftrightarrow_\eta \partial \Lambda \\ 0 & \text{if } v_0 \in \Lambda^3 \end{cases} \]

where $v \leftrightarrow_\eta \partial \Lambda$ denotes the event that $v$ is connected to $\partial \Lambda$ through a path of edges with $\eta_e = 1$ [note that $v_0 \in \Lambda^3$ implies $v_0 \not\leftrightarrow_\eta \partial \Lambda$]. For the unconditional law, it follows that

\[(5.7) \quad \rho^1_{\Lambda, \beta}(\sigma_{v_0} = 1) - \rho^1_{\Lambda, \beta}(\sigma_{v_0} = 2) = \rho^1_{\Lambda, \beta}(v_0 \leftrightarrow_\eta \partial \Lambda) . \]

For $v_1 \in V_1 \cap \Lambda$, one has similar equations with the roles of colors 1 and 2 interchanged, so that

\[(5.8) \quad \rho^1_{\Lambda, \beta}(\sigma_{v_1} = 2) - \rho^1_{\Lambda, \beta}(\sigma_{v_1} = 1) = \rho^1_{\Lambda, \beta}(v_1 \leftrightarrow_\eta \partial \Lambda) . \]

Now consider a finite set $\Delta_0 \subset V_0$, and recall that $J_{k \Delta_0}$ denotes the event that $\Delta_0$ is uniformly colored in the color $k$, and that $J_{\Delta_0} = \bigcup_{k=1}^3 J_{k \Delta_0}$ denotes the event that all sites in $\Delta_0$ are uniformly colored in some color. Let $\Delta_0 \leftrightarrow_\eta \partial \Lambda$ denote the event that there is at least one site in $\Delta_0$ that is connected to $\partial \Lambda$ through a path of edges with $\eta_e = 1$. Since

\[(5.9) \quad \rho^1_{\Lambda, \beta}(J_{1 \Delta_0} \mid \Lambda^3, \eta) = \rho^1_{\Lambda, \beta}(J_{2 \Delta_0} \mid \Lambda^3, \eta) \quad \text{a.s. on} \quad \Delta_0 \not\leftrightarrow_\eta \partial \Lambda \]
and
\[ \rho_{\Lambda,\beta}(J_{2,\Delta_0} | \Lambda^3, \eta) = 0 \quad \text{a.s. on } \Delta_0 \leftrightarrow \eta \partial \Lambda, \]
we see that
\[ \rho_{\Lambda,\beta}(J_1,\Delta_0) - \rho_{\Lambda,\beta}(J_2,\Delta_0) = \rho_{\Lambda,\beta}(J_{\Delta_0} \cap \{\Delta_0 \leftrightarrow \eta \partial \Lambda\}). \]

Now, recall that a finite set \( \Delta \subset V \) is termed thick if there exists a nonempty finite subset \( \Delta_1 \subset V_1 \) that is connected in \( G_1 \) and such that \( \Delta = \{v \in V : d_G(v, \Delta_1) \leq 1\} \).

We therefore fix some thick set \( \Delta \subset V \) and define \( \Delta_0 := \Delta \cap V_0 \) (since \( G \) is bipartite we have \( \Delta_1 = \Delta \cap V_1 \)).

Comparing (5.7)/(5.8)/(5.11) and noting that \( \rho_{\Lambda,\beta} \) can be replaced by \( \mu_{\Lambda,\beta} \) on the left-hand sides, we see that Lemma 2.3 is implied by the following claim:

**Lemma 5.1. (Comparison lemma)** Fix \( \beta_0 > 0 \) and let \( \Delta \subset V \) be thick. Then there exists an \( \epsilon > 0 \) such that
\[ \rho_{\Lambda,\beta}(v \leftrightarrow \eta \partial \Lambda \text{ for all } v \in \Delta) \geq \epsilon \rho_{\Lambda,\beta}(J_{\Delta_0} \cap \{\Delta_0 \leftrightarrow \eta \partial \Lambda\}) \]
uniformly for all \( \beta \in [\beta_0, \infty) \) and all simply connected finite sets \( \Lambda \supset \Delta \) such that \( \partial \Lambda \subset V_0 \). In fact, we can choose \( \epsilon = 3^{-|\Delta_1|} (1 - e^{-\beta_0}) |E_\Delta| \) where \( E_\Delta = \{\{u, v\} \in E : u, v \in \Delta\} \).

**Proof.** The proof is by a finite-energy argument: that is, to each \( (\sigma, \eta) \in J_{\Delta_0} \cap \{\Delta_0 \leftrightarrow \eta \partial \Lambda\} \) we associate a \( (\sigma'', \eta'') \in \{v \leftrightarrow \eta \partial \Lambda \text{ for all } v \in \Delta\} \); we then compute a lower bound on the ratio of \( \rho_{\Lambda,\beta}(\sigma'', \eta'') \) to the total \( \rho_{\Lambda,\beta}(\sigma'', \eta'') \) weight of the configurations \( (\sigma, \eta) \) that map onto it. The construction is in two steps \( (\sigma, \eta) \mapsto (\sigma', \eta') \mapsto (\sigma'', \eta'') \).

In the first step we recolor all spins \( (\sigma_0)_{v \in \Delta_1} \) to \( \sigma_v = 2 \) (leaving all other variables as is). In the second step we set all bond variables \( (\eta_e)_{e \in E_\Delta} \) to \( \eta''_e = 1 \) (again leaving all other variables as is). Let us now compute a lower bound on the ratio of weights, as follows:

Since \( (\sigma, \eta) \in J_{\Delta_0} \cap \{\Delta_0 \leftrightarrow \eta \partial \Lambda\} \) and \( \partial \Lambda \) is colored 1, it follows that \( \sigma_v = 1 \) for all \( v \in \Delta_0 \). Since \( \Delta \) is thick, every vertex in \( \Delta_1 \) has all its neighbors in \( \Delta_0 \). Therefore we can recolor all sites in \( \Delta_1 \) with the color 2 without increase in energy, i.e.
\[ \rho_{\Lambda,\beta}(\sigma', \eta') \geq \rho_{\Lambda,\beta}(\sigma, \eta). \]

We lose a factor 3\(|\Delta_1|\) because 3\(|\Delta_1|\) configurations \( (\sigma', \eta') \) map onto the same configuration \( (\sigma, \eta') \).

We now have \( \sigma'_u \neq \sigma'_v \) for all \( \{u, v\} \in E_\Delta \). Therefore \( \rho_{\Lambda,\beta}(\sigma'', \eta'') \) is precisely \( p^{E_\Delta} \) times the total \( \rho_{\Lambda,\beta}(\sigma'', \eta'') \) weight of the configurations \( (\sigma', \eta') \) that map onto it, where
\[ p = 1 - e^{-\beta} \geq 1 - e^{-\beta_0}. \]

**Remark.** The ideas in this section — in particular, formulas (5.7), (5.8) and (5.11) — have an obvious generalization to the \( q \)-state Potts antiferromagnet for any \( q \geq 2 \) on any bipartite graph (not necessarily a plane quadrangulation), where

\[ \text{In fact, (5.12) would still hold (with a worse } \epsilon \text{) even if there were an energy cost associated to this operation, provided that this energy cost is uniformly bounded.} \]
we condition on \( q - 2 \) colors and use the random-cluster representation for the remaining 2 colors. Indeed, as in [24, Appendix A], we may consider an even more general situation: Suppose that the vertex set \( V \) is partitioned as \( V = V_0 \cup V_1 \); then we can consider a Potts model with antiferromagnetic interactions on edges connecting \( V_0 \) to \( V_1 \) and ferromagnetic interactions on edges \( V_0 - V_0 \) and \( V_1 - V_1 \).

**Appendix A. Some facts about infinite planar graphs**

The purpose of this appendix is to collect some facts about infinite graphs, and in particular about infinite graphs embedded in the plane, that will be needed in the main part of the paper. An excellent general introduction to the theory of infinite graphs can be found in [16, Chapter 8]; but we shall require here some further facts that are scattered throughout the recent research literature, plus a few that appear to be new.

**A.1. Basic facts and definitions.** Recall that a graph is a pair \( G = (V, E) \) consisting of a (not necessarily finite) vertex set \( V \) and edge set \( E \). Unless mentioned otherwise, when we say “graph” we will always mean a *simple* graph, i.e., an undirected graph that has no loops or multiple edges. Thus, the elements of \( E \) (the edges) are unordered sets \( \{v, w\} \) containing two distinct elements of \( V \). Two vertices \( v, w \in V \) are called *adjacent* if \( \{v, w\} \in E \). An edge \( e \) containing a vertex \( v \) is said to be *incident* to \( v \). The *degree* of a vertex \( v \in V \) is the number of edges incident to it. We say that \( G \) is *finite* (resp. *countable*) if both \( V \) and \( E \) are finite (resp. countable). We say that \( G \) is *locally finite* (resp. *locally countable*) if every vertex has finite (resp. countable) degree.

A graph \( G' = (V', E') \) such that \( V' \subseteq V \) and \( E' \subseteq E \) is called a *subgraph* of \( G = (V, E) \); we also say that \( G \) *contains* \( G' \). If \( E' \) contains all edges \( \{v, w\} \in E \) with \( v, w \in V' \), then \( G' \) is called the subgraph of \( G \) *induced* by \( V' \). Likewise, if \( V' = \{v \in V : \exists w \in V \text{ s.t. } \{v, w\} \in E'\} \), then we call \( G' \) the subgraph *induced* by \( E' \).

We will say that a graph \( G \) is *connected* if for every proper subset \( W \subset V \) (the word “proper” means that \( W \neq \emptyset, V \) there is an edge \( \{v, w\} \in E \) with \( v \in V \setminus W, w \in W \). Let us note that every locally countable, connected graph is countable. \(^{22} \) A connected graph in which each vertex has degree \( \leq 2 \) will be called a *generalized path*. The *length* of a generalized path is the number of its edges. Vertices of degree 2 are called *internal vertices* of the generalized path, while vertices of degree one or zero are called *endvertices*. \(^{23} \) An infinite generalized path with one endvertex is called a one-way-infinite path or *ray*; an infinite generalized path without endvertices is called a two-way-infinite path or *double ray*; a finite generalized path without endvertices is called a *cycle*; and a finite generalized path

---

\(^{22}\) Indeed, if \( G = (V, E) \) is connected and \( W \) is the set of all vertices at finite graph distance from a given vertex, then connectedness implies \( W = V \).

\(^{23}\) Note that a vertex of degree zero can occur in a connected graph (and more particularly in a path) only if the graph has precisely one vertex.
with one or two endvertices is called a path. In particular, a graph consisting of a single vertex and no edges is a path of length zero.

Two vertices \( v, w \) in a graph \( G \) are linked by a path if \( G \) contains a path that has \( v \) and \( w \) as its endpoints. Then it is easy to see that a graph is connected (according to our definition above) if and only if every pair of vertices in \( G \) is linked by a path. The graph distance \( d(v, w) \) between two vertices \( v, w \in V \) is the length of a shortest path linking \( v \) and \( w \) if such a path exists, and \( \infty \) otherwise. An edge \( \{v, w\} \) in a path \( P \) is said to be a final edge of \( P \) if either \( v \) or \( w \) (or both) has degree one.

The graph distance \( d(e, f) \) between two edges \( e, f \in E \) is defined as the minimal length minus one of a path that has \( e \) and \( f \) as final edges. Note, in particular, that \( d(e, f) = 0 \) iff \( e = f \), and that \( d(e, f) = 1 \) iff \( e \neq f \) but \( e \) and \( f \) share a vertex. It is easy to check that, whenever \( G \) is connected, the graph distance between vertices (resp. edges) defines a metric on \( V \) (resp. \( E \)).

Let \( G = (V, E) \) be a connected graph. A set \( B \subseteq E \) is called separating if \( G \setminus B \) is not connected. A set \( C \subseteq E \) is called a cutset if there exists a partition \( \{V_1, V_2\} \) of \( V \) into two nonempty sets such that \( C = E(V_1, V_2) := \{\{v_1, v_2\} \in E: v_1 \in V_1, v_2 \in V_2\} \). (Since \( G \) is assumed connected, \( \emptyset \) is not a cutset.) A separating set (resp. cutset) is called minimal if it contains no proper subset that is a separating set (resp. a cutset). In fact, the two concepts are the same: each minimal separating set is also a minimal cutset and vice versa. Moreover, a set \( C \subseteq E \) is a minimal cutset if and only if \( (V, E \setminus C) \) has exactly two connected components.

Two rays in an infinite graph \( G \) are said to be end-equivalent (or equivalent for short) if one (hence all) of the following equivalent conditions holds:

1) there exists a third ray whose intersection with both of them is infinite;
2) there are infinitely many disjoint paths in \( G \) joining the two rays;
3) for every finite set \( S \subset V \), the two rays are eventually contained in the same connected component of \( G \setminus S \).

It is easy to see that end-equivalence is an equivalence relation. The corresponding equivalence classes are termed the ends of the graph \( G \). It is not hard to see that if \( G \) is connected, locally finite graph \( G = (V, E) \) has one end if and only if for every finite \( E' \subset E \), the subgraph \( (V, E \setminus E') \) has exactly one infinite component; or equivalently, if every finite minimal cutset divides \( V \) into two connected components, of which exactly one is of infinite size.

---

24 In the graph-theory literature, minimal cutsets are sometimes called bonds. Alas, in the statistical-mechanics literature, “bond” is often used as a synonym for “edge”. To forestall confusion, we prefer to avoid the term “bond” altogether.

25 This definition of the ends of a graph is due to Halin [34] in 1964 (see also Freudenthal [26] for the locally finite case). The notion of “end” of a topological space was introduced much earlier by Freudenthal [25], in 1931. If one identifies \( G \) with the “topological realization” defined below (at the beginning of Section A.3), then the two notions coincide for locally finite graphs but are different in general [18]. See [48] for an elementary introduction to the theory of ends of locally finite graphs; and see [50] for a survey treating both graphs and topological spaces.
For $k \geq 1$, a graph $G = (V, E)$ is called \textit{k-connected} if $|V| \geq k+1$ and the subgraph induced by $V \setminus W$ is connected for all $W \subset V$ satisfying $|W| < k$. (Otherwise put, to disconnect $G$ one must remove at least $k$ vertices.) Two vertices $v, w$ in $G$ are said to be \textit{k-edge-connected} if one needs to remove at least $k$ edges to unlink them; a graph is called \textit{k-edge-connected} if every pair of vertices in it is $k$-edge-connected. It is easy to see that $k$-edge-connectedness of vertices is an equivalence relation; the corresponding equivalence classes of vertices (and their induced subgraphs) are called the \textit{k-edge-connected components} of the graph. Two paths are called \textit{vertex-disjoint} (resp. \textit{edge-disjoint}) if their sets of internal vertices (resp. edges) are disjoint. By Menger’s theorem, two vertices are $k$-edge-connected if and only if they are linked by $k$ edge-disjoint paths, and a graph is called $k$-connected if and only if every two vertices are linked by $k$ vertex-disjoint paths.

An \textit{automorphism} of a graph $G = (V, E)$ is a bijection $g: V \to V$ such that $\{g(v), g(v')\} \in E$ if and only if $\{v, v'\} \in E$. We say that two vertices $v, w \in V$ are of the same \textit{type}, denoted $v \sim w$, if there exists an automorphism $g$ of $G$ such that $g(v) = w$. Then $\sim$ is an equivalence relation that divides the vertex set $V$ into equivalence classes called types. A graph is called \textit{vertex-transitive} if there is only one type of vertex, and \textit{vertex-quasi-transitive} if there are only finitely many types of vertices. Similarly, we say that two edges $\{v, v'\}, \{w, w'\} \in E$ are of the same type if there exists an automorphism $g$ of $G$ such that $\{g(v), g(v')\} = \{w, w'\}$; and we say that two directed edges $(v, v'), (w, w') \in V \times V$ with $\{v, v'\}, \{w, w'\} \in E$ are of the same type if there exists an automorphism $g$ of $G$ such that $g(v) = w$ and $g(v') = w'$. Edge- and directed-edge- transitivity or quasi-transitivity are then defined in the obvious way. We shall need the following fairly easy result:

\textbf{Lemma A.1.} For a locally finite graph $G$, the following are equivalent:

(a) $G$ is vertex-quasi-transitive.
(b) $G$ is edge-quasi-transitive.
(c) $G$ is directed-edge-quasi-transitive.

\textbf{Proof.} (b) \Leftrightarrow (c): Obviously, if two directed edges $(v, v')$ and $(w, w')$ are of the same type, then the corresponding undirected edges $\{v, v'\}$ and $\{w, w'\}$ are also of the same type. This shows that there are as most as many types of edges as there are types of directed edges. Conversely, since there are only two ways to order a set with two elements, there are at most twice as many types of directed edges as there are types of edges.

(c) \Rightarrow (a): If two directed edges $(v, v')$ and $(w, w')$ are of the same type, then obviously $v$ and $w$ are of the same type. Since all isolated vertices (i.e., vertices of degree zero) are of the same type, this shows that the number of types of vertices is at most the number of types of directed edges plus one.

(a) \Rightarrow (c): Assume that there are $m$ types of vertices and that these have degrees $d_1, \ldots, d_m$. Pick representatives $v_1, \ldots, v_m$ of these equivalence classes. For any directed edge $(v, v')$, there exists a $k \in \{1, \ldots, m\}$ and a graph automorphism that
maps $v$ to $v_k$. Since a graph automorphism preserves the graph structure, $w'$ must be mapped into one of the $d_k$ vertices adjacent to $v_k$. Thus, we have found $d_1 + \cdots + d_m$ directed edges such that each directed edge can be mapped into one of these by a graph automorphism. In particular, the number of types of directed edges is at most $d_1 + \cdots + d_m$. □

In view of Lemma A.1, we usually talk about quasi-transitive graphs without specifying whether we mean in the vertex, edge or directed-edge sense.²⁶

A geodesic in a graph $G$ is a generalized path $P$ such that for each pair of vertices $v, w$ in $P$, the graph distance from $v$ to $w$ in $P$ coincides with the graph distance from $v$ to $w$ in $G$. It is not hard to see that any path of minimal length linking two vertices $v', w'$ is a geodesic. For completeness, we prove the following simple fact. Part (a) of this lemma can be found, for example, in [59, Prop. 1]; it is a simple corollary of König’s Infinity Lemma [16, Lemma 8.1.2]. We did not find a reference for part (b), but it is presumably well known.

Lemma A.2. (Infinite geodesics) Let $G = (V, E)$ be a locally finite connected graph with infinite vertex set $V$. Then:

(a) For each $v \in V$, there exists a geodesic ray whose endpoint is $v$.

(b) If $G$ is moreover quasi-transitive, then $G$ contains a geodesic double ray.

Proof. Since each vertex is of finite degree, the set of vertices at distance $k$ from $v$ is finite for each $k \geq 0$. Therefore, since $V$ is infinite and $G$ is connected, for each $n \geq 1$ we can find a vertex $v_n \in V$ at distance $d(v, v_n) = n$. Consider a path $(v_0^{(n)} = v, v_1^{(n)}, \ldots, v_{n-1}^{(n)}, v_n^{(n)} = v_n)$ and define the function $f_n: \mathbb{N} \to V$ by taking $f_n(k) = v_k^{(n)}$ for $k \leq n$ and $f_n(k) = v_n$ for all $k \geq n$. Since the set of points at distance $k$ from $v$ is finite for each $k \geq 0$, we may select a subsequence $f_{n_m}$ that converges pointwise in the discrete topology. It is easy to see that the limit of such a subsequence is a geodesic ray starting at $v$, proving part (a) of the lemma.

To prove also part (b), we use that by quasi-transitivity, the geodesic ray constructed in part (a) must pass through at least one type of vertex infinitely often. It follows that for a vertex $v$ of this type, for each $n \geq 0$ we can find a function $f_n: \mathbb{Z} \to V$ such that $f_n(0) = v$, $f_n(k) = f_n(-n)$ for all $k \leq -n$, the vertices \{ $f_n(k): k \geq -n$ \} are all different, and the graph induced by this set is a geodesic ray. (It suffices to employ the corresponding automorphism to shift the original geodesic ray by identifying a vertex of given type sufficiently far on it with the vertex $v$.) Now the statement follows from the same sort of compactness argument as used in the proof of part (a). □

²⁶ The term “almost transitive” is also used as a synonym for “quasi-transitive”.
A.2. Duality. In the main part of this paper, we make extensive use of the fact that the two sublattices \( G_0 \) and \( G_1 \) are each other’s dual in the sense of planar graph duality. Such duals may be defined abstractly, using only basic concepts of graph theory, without any reference to embeddings of graphs in the plane. In fact, this abstract notion of duality is sufficient for all our proofs, as we shall show in the present subsection. Nevertheless, in the next subsection we will complement this abstract theory by showing that sufficiently “nice” embeddings of a graph in the plane give rise to abstract duals, and conversely that every locally finite abstract dual arises in this way.

The basic theory of duals of infinite graphs was developed by Thomassen [68, 69], which we largely follow here; see also Bruhn and Diestel [9, 17, 10] for a partially alternative approach. Abstract duals can be defined for 2-connected graphs, but in this case the dual may have multiple edges; we therefore restrict ourselves for simplicity to the 3-connected case.

If \( G = (V, E) \) is a 3-connected graph, then an abstract dual of \( G \) is a connected graph \( G^\dagger = (V^\dagger, E^\dagger) \) together with a bijection \( E \ni e \mapsto e^\dagger \in E^\dagger \) such that a finite set \( C \subseteq E \) is a cycle in \( G \) if and only if \( C^\dagger := \{e^\dagger : e \in E\} \) is a minimal cutset in \( G^\dagger \). We stress that in this generality the term “dual” is something of a misnomer, since \( G \) need not be an abstract dual of \( G^\dagger \); indeed, \( G^\dagger \) need not have any abstract dual at all, even when \( G \) is planar and locally finite [68, p. 266]. However, if both \( G \) and \( G^\dagger \) are locally finite, then the situation becomes particularly nice:

**Theorem A.3. (Locally finite abstract duals)** Let \( G = (V, E) \) be a locally finite 3-connected graph that has a locally finite abstract dual \( G^\dagger = (V^\dagger, E^\dagger) \). Then:

(i) \( G \), with the inverse map \( E^\dagger \ni e^\dagger \mapsto e \in E \), is an abstract dual of \( G^\dagger \).

(ii) \( G^\dagger \) is 3-connected.

(iii) \( G^\dagger \) is, up to isomorphism, the only abstract dual of \( G \).

(iv) If \( G \) has one end, then so does \( G^\dagger \).

(v) If \( G \) is quasi-transitive, then so is \( G^\dagger \).

**Proof.** Part (ii) follows from [69, Theorem 4.5]. Then parts (i) and (iii) follow from [68, Theorem 9.5]. Part (iv) follows from [10, Theorem 1.1], which states that there is a homeomorphism between the spaces of ends of \( G \) and \( G^\dagger \) (considered as subspaces of the Freudenthal compactification, to be defined in the next subsection); so in particular \( G \) has one end if and only if \( G^\dagger \) does. (Alternatively, this can be deduced from Proposition A.12(iii) below, using Lemma A.7 and Theorem A.10.)

We did not find a reference for part (v), but this is not hard to prove using some more results from [68]. We will prove the following, stronger statement. Let

\footnote{The approach of Thomassen [68, 69] is based on the study of finite cycles and minimal cutsets, as explained below. The alternative approach of Bruhn and Diestel [9, 17, 10] introduces (by topological means) a notion of infinite “cycles”, and shows that this notion allows a somewhat cleaner duality theory. When \( G^\dagger \) is locally finite, the two concepts of duality coincide [9, Lemma 4.7]. We are therefore entitled to use here the theorems from [9, 17, 10] under the added hypothesis that \( G^\dagger \) is locally finite.}
Let $G$ and $G^\dagger$ be as in Theorem [A.3], let $v^\dagger \in V^\dagger$ be a vertex in $G^\dagger$, and let $E^\dagger_{v^\dagger} := \{e^\dagger \in E^\dagger: e^\dagger \text{ is incident to } v^\dagger\}$. Then $E^\dagger_{v^\dagger}$ is a minimal cutset in $G^\dagger$, which is finite by virtue of the local finiteness of $G^\dagger$; hence $E^\dagger_{v^\dagger} := \{e^\dagger \in E^\dagger: e^\dagger \in E^\dagger_{v^\dagger}\}$ is a cycle in $G$. We loosely call $v^\dagger$ a *face* of $G$ and we call $E^\dagger_{v^\dagger}$ the *boundary* of this face. Indeed, we will see in the next subsection that for a suitable embedding of $G$ in the plane $\mathbb{R}^2$, $v^\dagger$ corresponds to a connected component of $\mathbb{R}^2 \setminus G$ and $E^\dagger_{v^\dagger}$ corresponds to its boundary. In view of this, we abstractly define a *triangulation* (resp. *quadrangulation*) to be a locally finite 3-connected graph $G$ that has an abstract dual $G^\dagger$ in which each vertex has degree 3 (resp. 4).

Now assume that $G$ has one end. Then each finite minimal cutset $C$ of $G$ corresponds to a partition $\{V_1, V_2\}$ of $V$ into two connected components, of which exactly one is infinite. Let $V_1, V_2$ denote the finite and infinite component, respectively. Since by Theorem [A.3(i)], $G$ is an abstract dual of $G^\dagger$, the set $C^\dagger := \{e^\dagger: e \in C\}$ is a cycle in $G^\dagger$, and each cycle in $G^\dagger$ arises in this way. We call $\text{Int}(C^\dagger) := V_1$
and \( \text{Ext}(C^\dagger) := V_2 \) the interior and exterior of \( C^\dagger \), respectively. We say that \( C^\dagger \) surrounds a vertex \( v \in V \) if \( v \in \text{Int}(C^\dagger) \).

An essential ingredient in our proofs in this paper is an upper bound \( (\text{Lemma 3.4}) \) for certain quasi-transitive triangulations \( G \) on the number of cycles in \( G^\dagger \) of a given length surrounding a given vertex \( v \) in \( G \). To derive this bound, we need some simple graph-theoretic facts.

**Lemma A.4. (Distances in a graph and its dual)** Let \( G = (V,E) \) be a 3-connected graph. Assume that each vertex in \( G \) has a locally finite abstract dual \( G^\dagger \). Then for all \( e,f \in E \) we have

\[
\begin{equation}
\tag{A.1}
d(e^\dagger,f^\dagger) \leq \left( \frac{1}{2}d_{\max} - 1 \right) d(e,f) + 1,
\end{equation}
\]

where \( d(e^\dagger,f^\dagger) \) denotes the distance between \( e^\dagger \) and \( f^\dagger \) in the dual graph \( G^\dagger \).

Note that \( d_{\max} \geq 3 \) since \( G \) is 3-connected.

**Proof of Lemma A.4.** If \( e = f \) (hence \( e^\dagger = f^\dagger \)), the statement is trivial, so consider the case \( e \neq f \). Let \( P \) be a path of minimal length that has \( e \) and \( f \) as final edges. For each vertex \( v \) of \( G \), let \( E_v^\dagger := \{ e^\dagger : e \text{ is incident to } v \} \) denote the boundary of the corresponding face of \( G^\dagger \). Note that \( E_v^\dagger \) is a cycle whose length is the degree of \( v \). If some \( E_v^\dagger \) and \( E_w^\dagger \) share an edge \( e^\dagger \), then \( e \) connects \( v \) and \( w \). For internal vertices \( v,w \) of \( P \), by the minimality of \( P \), this is possible only if \( v \) and \( w \) are adjacent in \( P \). Let \( v_1,\ldots,v_n \) be the internal vertices of \( P \), where \( n = d(e,f) \), and let \( d_1,\ldots,d_n \) denote their degrees in \( G \). Then the symmetric difference \( D := E_v^\dagger \setminus E_w^\dagger \) consists of exactly \( \sum_{k=1}^n d_k - 2(n-1) \) edges which form a cycle in \( G^\dagger \) containing \( e^\dagger \) and \( f^\dagger \). It follows that \( G^\dagger \) contains two paths \( P_1^\dagger, P_2^\dagger \) which have \( e^\dagger \) and \( f^\dagger \) as their final edges and are otherwise edge-disjoint, and whose respective lengths \( k_1,k_2 \) satisfy \( k_1 + k_2 - 2 = \sum_{k=1}^n d_k - 2(n-1) \). We conclude from this that \( \min(k_1,k_2) \leq (k_1 + k_2)/2 \leq \frac{1}{2}(nd_{\max} - 2n + 4) \) and hence \( d(e^\dagger,f^\dagger) \leq \left( \frac{1}{2}d_{\max} - 1 \right)n + 1 \).

Now let \( G = (V,E) \) be a locally finite 3-connected quasi-transitive graph that has a locally finite abstract dual \( G^\dagger = (V^\dagger,E^\dagger) \). By Theorem A.3(v), \( G^\dagger \) is also quasi-transitive, hence of bounded degree; let \( d_{\max}^\dagger \) denote the maximal degree of a vertex in \( G^\dagger \). By Lemma A.2(b), \( G \) contains at least one geodesic double ray; let \( V_0 \) be the set of all vertices \( v \in V \) that lie on some geodesic double ray. Then, by quasi-transitivity, the maximal distance of any vertex in \( G \) to the set \( V_0 \),

\[
\begin{equation}
\tag{A.2}
K := \sup_{w \in V} \inf_{v \in V_0} d(v,w),
\end{equation}
\]

is finite. Note, finally, that by Lemma A.2(a), at each \( v \in V \) there starts at least one geodesic ray.

**Proposition A.5. (Distance to a surrounding cycle)** Let \( G, G^\dagger, d_{\max}^\dagger \) and \( K \) be as above. Assume that \( G \) has one end. Let \( v \in V \), let \( R \) be a geodesic ray in \( G \)
starting at \(v\), and let \(C^\dagger\) be a cycle in \(G^\dagger\) of length \(L\) surrounding \(v\). Then \(C^\dagger\) must cross one of the first \(N\) edges of \(R\), where

\[
N := 1 + K + \frac{1}{2}(d_{\text{max}}^\dagger - 1) L.
\]

**Proof.** As discussed above, the cycle \(C^\dagger\) corresponds to a minimal cutset \(C\) of \(G\), which divides \(G\) into two connected components, one of which is finite and the other of which is infinite; moreover, \(v\) is contained by hypothesis in the finite component. Since \(R\) is an infinite ray starting at \(v\), it must use somewhere an edge of \(C\); let \(f\) be the first such edge, and let \(f^\dagger \in C^\dagger\) be the corresponding dual edge.

Let \(w\) be the point in \(V_0\) that is closest to \(v\), let \(P\) be a path of minimal length \(K \leq K\) linking \(v\) and \(w\), and let \(D\) be a geodesic double ray containing \(w\). Write \(D = R_1 \cup R_2\) where \(R_1, R_2\) are geodesic rays starting at \(w\), and observe that \(R_1^\dagger := P \cup R_1\) and \(R_2^\dagger := P \cup R_2\) are rays starting at \(v\). The cycle \(C^\dagger\) must cross some edge in \(R_1^\dagger\) and some edge in \(R_2^\dagger\). Let \(e_1, e_2\) be the first edges (counting from \(v\)) in \(R_1^\dagger, R_2^\dagger\) crossed by \(C^\dagger\). We distinguish two cases: I. \(e_1 \neq e_2\) and II. \(e_1 = e_2\).

In case I, \(e_1\) and \(e_2\) lie on \(D\) and are the \((K + N_1)\)-th and \((K + N_2)\)-th edge of the rays \(R_1^\dagger\) and \(R_2^\dagger\), say. Then \(C^\dagger\) is the union of two paths \(P_1^\dagger, P_2^\dagger\), each of which has \(e_1^\dagger\) and \(e_2^\dagger\) as their final edges, and which are disjoint except for their overlap at \(e_1^\dagger\) and \(e_2^\dagger\). Let \(L_1, L_2\) denote the lengths of these paths in \(G^\dagger\), where \(L_1 + L_2 - 2 = L\).

Without loss of generality we may assume that \(f^\dagger\) lies on \(P_1^\dagger\) and is the \(M_1\)-th edge of \(P_1^\dagger\) starting from \(e_1\) and the \(M_2\)-th edge starting from \(e_2\), where \(M_1 + M_2 - 1 = L_1\). By Lemma A.4 (applied with the roles of \(G\) and \(G^\dagger\) reversed), we have

\[
\min(L_1, L_2) \geq d(e_1^\dagger, e_2^\dagger) + 1 \geq \frac{1}{c}(d(e_1, e_2) - 1) + 1 = \frac{1}{c}(N_1 + N_2 - 2) + 1,
\]

where we have abbreviated \(c := \frac{1}{2}d_{\text{max}}^\dagger - 1\). Therefore

\[
M_1 + M_2 = L_1 + 1 = L + 3 - L_2 \leq L + 2 - \frac{1}{c}(N_1 + N_2 - 2).
\]

Using Lemma A.4 again, there exists a path in \(G\) of length at most \(c(M_1 - 1) + 1\) that has \(e_1\) and \(f\) as its final edges, and another path of length at most \(c(M_2 - 1) + 1\) that has \(e_2\) and \(f\) as its final edges. Combining these paths with the pieces of \(R_1^\dagger\) and \(R_2^\dagger\) leading up to \(e_1\) and \(e_2\), respectively, we find two paths in \(G\) starting at \(v\) and with \(f\) as their final edge, with lengths of at most

\[
K + N_1 + c(M_1 - 1) \quad \text{and} \quad K + N_2 + c(M_2 - 1),
\]

respectively. By (A.5), it follows that the average length of these two paths is at most

\[
\frac{1}{2}(2K + N_1 + N_2 + c(M_1 + M_2 - 2)) \leq \frac{1}{2}
\]

\[
\left(2K + N_1 + N_2 + c(L - \frac{1}{c}(N_1 + N_2 - 2))\right) = 1 + K + \frac{1}{2}cL.
\]
Taking the shorter of these two paths, we have found a path of length at most 
\[ 1 + K + \frac{1}{2}cL \]
starting at \( v \) and having \( f \) as its final edge. But since \( R \) is a geodesic 
ray, the distance from \( v \) to \( f \) along \( R \) must be at most this.

In case II, \( e := e_1 = e_2 \) must lie on \( P \); it is the first edge of \( P \) (counting from \( v \)) 
that is crossed by some edge in \( C^\dagger \). Then \( C^\dagger \) contains two paths with lengths \( L_1, L_2 \) 
satisfying \( L_1 + L_2 - 2 = L \) that have \( e^\dagger \) and \( f^\dagger \) as their final edges. It follows that 
\[ d(e^\dagger, f^\dagger) \leq \frac{1}{2}L. \]
Therefore, by Lemma A.4, \( d(e, f) \leq \frac{1}{2}cL + 1 \), which means that we can 
find a path in \( G \) of length at most \( \frac{1}{2}cL + 2 \) that has \( e \) and \( f \) as its final edges.

Combining this path with the piece of \( P \) leading from \( v \) to \( e \), we again find a path 
of length at most \( 1 + K + \frac{1}{2}cL \) starting at \( v \) and having \( f \) as its final edge. \( \square \)

A.3. Planar embeddings. In this section we collect some results that show that 
sufficiently “nice” embeddings of a graph in the plane give rise to a geometric dual 
that is also an abstract dual, and conversely that graphs having a locally finite 
abstract dual have “nice” embeddings in the plane such that the geometric dual 
coincides with the abstract dual.

Embeddings of finite (planar) graphs in the plane are treated in almost any ele-
mentary book on graph theory, but it is more difficult to find a good reference for 
infinite planar graphs. Some articles that we have found useful are [6, 9, 19, 20, 49, 
64, 68, 69].

Each graph \( G \) gives rise to a topological space — which, by a slight abuse of 
notation, we shall continue to call \( G \) — that is defined by first assigning a disjoint 
copy of \( [0, 1] \) to each edge of \( G \) and a point to each vertex of \( G \) and then identifying 
the endpoints of intervals with the endvertices of the corresponding edges.

We equip \( G \) with the quotient topology arising from this identification: thus, a neighbor-
hood base of an inner point on an edge is formed by the open intervals on the 
edge containing that point, while a neighborhood base of a vertex \( x \) is formed by the 
unions of half-open intervals \( [x, z) \) containing \( x \), one interval being taken from every 
edge \( [x, y] \) incident to \( x \).\([28]\) Such a topological realization of the graph \( G \) is compact 
if and only if \( G \) is finite, and locally compact if and only if \( G \) is locally finite. It is 
metrizable if and only if \( G \) is locally finite.\([30]\)

---
\( ^{28} \) More precisely, given a graph \( G = (V, E) \), we start from the set \((E \times [0, 1]) \cup (V \times \{2\})\), and then for each edge \( e = xy \) we identify \((e, 0)\) with \((x, 2)\) and \((e, 1)\) with \((y, 2)\).

\( ^{29} \) With this topology, \( G \) is a 1-dimensional CW-complex [35] pp. 5–6 and pp. 519 ff.]

\( ^{30} \) Indeed, if \( x \) is a vertex of infinite degree, then \( x \) does not have a countable neighborhood 
basis in the topology we have given \( G \). (It is possible to equip \( G \) with a different topology, which is 
always metrizable, in such a way that a neighborhood base of a vertex \( x \) is formed by the unions of 
half-open intervals \( [x, z) \) using the same distance \( \epsilon = d(x, z) \) for each incident edge \([28]\). However, 
we shall not use this topology.)
An **embedding** of a graph $G$ in the plane is a continuous injective map $\phi: G \to \mathbb{R}^2$.

A graph $G$ that can be embedded in the plane is called **planar**. A **plane graph** is a pair $(G, \phi)$ where $G$ is a graph and $\phi$ is an embedding of $G$ in the plane. We (topologically) identify the sphere $\mathbb{S}$ with the one-point compactification $\mathbb{R}^2 \cup \{\infty\}$ of the plane $\mathbb{R}^2$. Embeddings of graphs in the sphere are defined analogously to embeddings in the plane; a graph can be embedded in the sphere if and only if it can be embedded in the plane.

If $G$ is a finite graph, then $\phi(G)$, being the continuous image of a compact set, is a closed subset of $\mathbb{R}^2$. Moreover, $\phi$ is necessarily a homeomorphism to its image, i.e., the inverse map $\phi^{-1}: \phi(G) \to G$ is also continuous. Both statements fail in general when $G$ is infinite. Mainly for these reasons, not much can be said in general about embeddings of infinite graphs; one needs extra conditions to proceed.

So let $G$ be a graph with no isolated vertices, and let $\phi: G \to \mathbb{S}$ be an embedding of $G$ in the sphere. Following an idea of [49], we say that $\phi$ is **self-accumulation-free** if no point $z \in \phi(G)$ is an accumulation point of edges that do not contain $z$. We say that $\phi$ is **pointed** if the image under $\phi$ of each ray in $G$ converges to a point in $\mathbb{S}$ (Clearly, equivalent rays must converge to the same point. Inequivalent rays may or may not converge to the same point.)

Recall that a compactification $\overline{F}$ of a topological space $F$ is a compact topological space $\overline{F}$ such that $F \subseteq \overline{F}$ is dense. We will always require that $\overline{F}$ be Hausdorff. As in [64], we say that a compactification $\overline{G}$ of a connected, locally finite (not necessarily planar) graph $G$ is **pointed** if each ray in $G$ converges to some point in $\overline{G}$ (Clearly, equivalent rays must converge to the same point. Inequivalent rays may or may not converge to the same point.) If $G$ is a graph and $\overline{G}$ is any compactification...
of $G$, then an embedding of $G$ in the sphere is a continuous injective map $\overline{\phi} : \overline{G} \to \mathbb{S}$; since $G$ is compact, it is necessarily a homeomorphism to its image.

**Lemma A.6. (Embeddings of compactifications)** Let $(G, \phi)$ be a locally finite plane graph with no isolated vertices. Then the following conditions are equivalent:

(i) $\phi$ is self-accumulation-free.

(ii) $\phi$ is a homeomorphism to its image.

(iii) $\phi$ can be extended to an embedding $\overline{\phi} : \overline{G} \to \mathbb{S}$ in the sphere of some compactification $\overline{G}$ of $G$. [The extension is of course unique, since $G$ is dense in $\overline{G}$.] Moreover, under these conditions, the embedding $\phi$ determines the compactification $\overline{G}$ uniquely (up to trivial renamings of points in $\overline{G}\setminus G$); and $\overline{G}$ is a pointed compactification if and only if $\phi$ is a pointed embedding.

**Proof.** If $\phi$ is not a homeomorphism to its image, then we can find $x_n, x \in G$ such that $\phi(x_n) \to \phi(x)$ but $x_n \not\to x$. Since $\phi$ is continuous and injective, the sequence $(x_n)$ cannot have any accumulation points in $G$ other than $x$; so by passing to a subsequence way may assume that it has no accumulation points at all. Since any finite collection of edges is compact, the sequence $(x_n)$ must visit infinitely many edges. Since $G$ is locally finite, only finitely many of these edges can contain $x$. It follows that $\phi$ is not self-accumulation-free.

Conversely, if $\phi$ is not self-accumulation-free, then there exists $z = \phi(x) \in \phi(G)$ and a sequence $(x_n)$ belonging to edges not containing $x$, such that $\phi(x_n) \to \phi(x)$. Clearly $x_n \not\to x$, so $\phi$ is not a homeomorphism to its image.

If (iii) holds, then $\overline{\phi}$ is a homeomorphism to its image, hence so is its restriction $\phi$ to $G$. Conversely, if $\phi$ is a homeomorphism to its image, then we may topologically identify $G$ with its image $\phi(G)$. Then the closure $\overline{\phi(G)}$ of $\phi(G)$ in the sphere $\mathbb{S}$ is a compactification of $\phi(G)$ such that the identity map from $\phi(G)$ to itself can be continuously extended to $\overline{\phi(G)}$; moreover, $\overline{\phi(G)}$ is (up to renaming) the only compactification of $\phi(G)$ with this property.

Clearly, if $\overline{\phi(G)}$ is a pointed compactification, then $\phi$ is a pointed embedding. Conversely, if $\phi$ is a pointed embedding, then each ray in $G$ converges to some point $z \in \mathbb{S}$; but since $\phi$ is self-accumulation-free, we have $z \notin \phi(G)$, hence $\overline{\phi(G)}$ is a pointed compactification. □

**Remark.** If $G$ is not locally finite, then $\phi$ can never be a homeomorphism of $G$ (with the topology we have given it) to its image. For if $x$ is a vertex of infinite degree, then it is easy to choose points $x_n$ lying on infinitely many distinct edges such that $\phi(x_n) \to \phi(x)$; but $x_n \not\to x$.

Let $\phi : G \to \mathbb{S}$ be any embedding of a graph $G$ in the sphere. Following [68, 69], let us define a vertex accumulation point (resp. edge accumulation point) of an embedding $\phi$ to be a point in $\mathbb{S}$ such that each of its open neighborhoods contains infinitely
many vertices (resp. intersects infinitely many edges). We abbreviate “vertex accumulation point” and “edge accumulation point” by VAP and EAP, respectively. Let us also say that a point \( x \in S \) is an endpoint of \((G, \phi)\) if there exists a ray in \( G \) such that its image under \( \phi \) converges to \( x \). The next lemma says that for self-accumulation-free pointed embeddings of 2-connected graphs, all these concepts coincide:

**Lemma A.7. (Endpoints of plane graphs)** Let \( G \) be a locally finite 2-connected graph, and let \( \phi: G \to S \) be a self-accumulation-free pointed embedding of \( G \). Then the following four sets are equal:

(i) The set of vertex accumulation points of \((G, \phi)\).

(ii) The set of edge accumulation points of \((G, \phi)\).

(iii) The set of endpoints of \((G, \phi)\).

(iv) \( \phi(G) \setminus \phi(G) \).

**Proof.** We will prove the inclusions (iii) \( \subseteq \) (i) \( \subseteq \) (iv) \( \subseteq \) (iii) and the same with (i) replaced by (ii). Clearly, every endpoint is also a vertex and edge accumulation point. Since \( \phi \) is self-accumulation-free and locally finite, each vertex or edge accumulation point lies in \( \phi(G) \setminus \phi(G) \). For each point \( x \in \phi(G) \setminus \phi(G) \), we can find a sequence of points \( x_n \in \phi(G) \), all lying on different edges, such that \( x_n \to x \). Since \( G \) is 2-connected, it can be shown that there exists a ray \( R \) in \( G \) whose image under \( \phi \) passes through infinitely many of these points. \( \square \) Since \( \phi \) is pointed, it follows that \( x \) is the limit of the \( \phi \)-image of \( R \).

Let \( \overline{G} \) be a pointed compactification of a connected locally finite (not necessarily planar) graph \( G \). Then, by definition, each ray in \( G \) converges to some limit in \( \overline{G} \setminus G \); and conversely, if \( G \) is 2-connected, then the argument proving (iv) \( \subseteq \) (iii) in Lemma A.7 shows more generally \( \text{[64, p. 4591]} \) that each point in \( \overline{G} \setminus G \) is the limit point of some ray in \( G \).

Each connected locally finite graph \( G \) has a unique (up to renaming) pointed compactification \( \overline{G} \) such that each point in \( \overline{G} \setminus G \) is the limit point of some ray in \( G \) and moreover two nonequivalent rays always converge to different limit points: this is the **Freudenthal compactification** \( \mathcal{F}(G) \), see \( \text{[16, Section 8.5]} \) or \( \text{[64, Section 7]} \). Clearly, \( \mathcal{F}(G) \setminus G \) is in bijection with the space of ends of \( G \). If \( G \) is 2-connected, then, in a sense, \( \mathcal{F}(G) \) is the “largest” pointed compactification of \( G \) (compare the remarks on \( \text{[64, p. 4594]} \)). At the other end of the scale, every connected, locally

---

35 Given an infinite set \( B \subseteq E \), we can use local finiteness to extract an infinite subset \( B' \subseteq B \) that is pairwise vertex-disjoint; then by \( \text{[64, Proposition 8]} \), \( G \) has a ray that uses infinitely many of the edges in \( B' \). Given an infinite subset \( T \subseteq V \), we can obviously choose an infinite set \( B \) of edges containing all the vertices in \( T \), and then proceed as before.

36 The Freudenthal compactification is a general construction of point-set topology: it is defined for locally compact Hausdorff spaces, or more generally for completely regular rim-compact spaces (see e.g. \( \text{[13, 20]} \)). For **locally finite** graphs — which are the only ones we are concerned with — this topological definition coincides with the graph-theoretic definition given here or in \( \text{[16, Section 8.5]} \).
finite, infinite graph $G$ has a smallest pointed compactification, namely the one-point compactification $G^* = G \cup \{\infty\}$ in which every infinite sequence of distinct vertices or edges (hence in particular every ray) converges to the single point $\infty$.

By Lemma A.6 an embedding $\phi$ of a connected locally finite graph $G$ is self-accumulation-free and pointed if and only if $\phi$ extends to an embedding $\overline{\phi}$ of a pointed compactification $\overline{G}$ of $G$. Richter and Thomassen [64, Theorems 1 and 13] have proven the following key result concerning the existence and uniqueness of such embeddings:

**Theorem A.8. (Embeddings of 3-connected planar graphs)** Let $G$ be a 3-connected locally finite planar graph. Then:

(a) There exists an embedding of the Freudenthal compactification $\mathcal{F}(G)$ in the sphere $S$.

(b) If $\overline{G}$ is any pointed compactification of $G$ and $\overline{\phi}_1, \overline{\phi}_2$ are embeddings of $\overline{G}$ in the sphere $S$, then there exists a homeomorphism $h: S \to S$ such that $\overline{\phi}_2 = h \circ \overline{\phi}_1$.

The unique (up to homeomorphism) embedding of the Freudenthal compactification $\mathcal{F}(G)$ in the sphere $S$ will be called the *Freudenthal embedding*.

Henceforth we assume that $G$ is a 2-connected locally finite graph and that $\phi: G \to S$ is a self-accumulation-free pointed embedding of $G$ in the sphere. By Lemma A.6 $\phi$ extends uniquely to an embedding $\overline{\phi}: \overline{G} \to S$ of an essentially unique pointed compactification $\overline{G}$ of $G$, and we have $\overline{\phi}(G) = \overline{\phi}(G)$. The connected components of the remaining open set $S \setminus \overline{\phi}(G)$ are called the *faces* of the plane graph $(G, \phi)$ [and also of $(\overline{G}, \overline{\phi})$]. For any face $f$, we let $\partial f := \overline{f} \setminus f$ denote its (topological) boundary. If $X$ is any topological space, we say that a subset $C \subseteq X$ is a *circle* if it is homeomorphic to the unit circle $S^1$. We need the following fundamental facts about the boundaries of faces:

**Theorem A.9. (Boundaries of faces)** Let $G$ be a 2-connected locally finite (planar) graph, and let $\phi$ be a self-accumulation-free pointed embedding of $G$ in the sphere $S$. Then each face $f$ of $(G, \phi)$ is bounded by a circle, and exactly one of the following three possibilities holds:

(i) $\partial f$ is the $\phi$-image of a (finite\footnote{Of course, cycles in the standard graph-theoretic meaning of the word are by definition finite. We will always use the word in this standard sense; we shall not make explicit use of the “infinite cycles” introduced (by topological means) by Diestel and collaborators [19, 20, 9, 17, 10].} cycle in $G$ [hence $\partial f \cap [\overline{\phi}(G) \setminus \phi(G)] = \varnothing$].

(ii) $\partial f \cap \phi(G)$ is the disjoint union of a nonempty countable collection of double rays, and $\partial f \cap [\overline{\phi}(G) \setminus \phi(G)] \neq \varnothing$.

(iii) $\partial f \cap \phi(G) = \varnothing$ [hence $\partial f \cap [\overline{\phi}(G) \setminus \phi(G)] = \partial f \simeq S^1$].

Moreover, when $\overline{G}$ is the Freudenthal compactification $\mathcal{F}(G)$, the set $\partial f \cap \phi(G)$ is dense in $\partial f$; in particular, case (iii) cannot occur.
Proof. By Lemma A.6, \( \phi \) extends uniquely to an embedding \( \overline{\phi} \) of an (essentially unique) pointed compactification \( \overline{G} \) of \( G \). Richter and Thomassen [64, Proposition 3 and Theorem 7] have proven the fundamental result that every face of \( (\overline{G}, \overline{\phi}) \) is bounded by a circle. By Lemma A.6, such a circle is the image of a circle \( C \) in \( \overline{G} \) under the homeomorphism \( \overline{\phi} \). We shall therefore prove the more general result that circles in \( \overline{G} \) have properties analogous to those stated in the theorem (here \( G \) need not be planar). We first claim that

(a) Whenever \( C \) contains an inner point of an edge \( e \in G \), it contains the entire edge \( e \).

(b) Whenever \( C \) contains a vertex \( x \in G \), it contains precisely two edges of \( G \) that are incident to \( x \).

For the Freudenthal compactification, (a) and (b) are proven in [20, Lemma 2.3]. But since statements (a) and (b) concern only a neighborhood of a point of \( G \), they hold for any Hausdorff compactification of \( G \) (since \( G \) is locally compact). For the Freudenthal compactification, it is moreover proven in [19, Lemma 4.3] that \( C \cap G \) is dense in \( C \).

By (a,b), the union of all edges belonging to \( C \) is a 2-regular graph \( C \cap G \), so each connected component of \( C \cap G \) is either a cycle or a double ray. Since \( C \cap G \) is moreover homeomorphic to a subset of \( S^1 \), we conclude that it is either empty [case (iii)], a cycle [case (i)], or a nonempty disjoint union of double rays [case (ii)]. In the latter case, there can be only countably many double rays (since \( G \) is countable), and \( C \) must also contain some points of \( \overline{G} \setminus G \) (in fact, at least as many as there are double rays) in order to have the topology of a circle. □

Examples. 1. The lattices shown in Figure 1 have one end, which is mapped to \( \infty \), and the boundary of every face is a cycle in \( G \). As we will see in Proposition A.16 below, quasi-transitivity and the fact that there is only one end imply that every face is bounded by a cycle.

2. Consider \( \mathbb{N} \times \mathbb{Z} \) with nearest-neighbor edges, with its usual embedding in the plane. This graph has one end, which is mapped to \( \infty \). It has a face that is bounded by a double ray together with \( \{ \infty \} = \overline{\phi(G)} \setminus \phi(G) \).

3. The graph in the top left of Figure 6 has uncountably many ends, which are mapped by the embedding \( \phi \) onto the unit circle. This embedding has one face that is bounded by endpoints alone (namely, the exterior face), as well as faces that are bounded by cycles and faces that are bounded by a double ray together with one endpoint. This graph is constructed in the following way. Consider a 3-regular tree with origin \( \varnothing \). Each vertex of \( G \) is labeled by a finite (possibly empty) sequence \( \ell_0 \ell_1 \ell_2 \cdots \ell_n \), where \( \ell_0 \in \{0, 1, 2\} \) and \( \ell_i \in \{0, 1\} \) for \( i \geq 1 \); and each ray emanating from \( \varnothing \) is labeled by an infinite sequence \( \ell_0 \ell_1 \ell_2 \cdots \) of the same type. The tree can
now be embedded in the unit disk such that the ray $\ell_0\ell_1\ell_2\cdots$ converges to the point
\begin{equation}
2\pi \cdot \frac{1}{3} \left( \ell_0 + \sum_{i=1}^{\infty} 2^{-i}\ell_i \right)
\end{equation}
on the unit circle. We then make the tree into a 3-connected graph by adding edges. For each vertex $\ell \neq \emptyset$, we draw edges between the points $\ell_{01}$ and $\ell_{10}$ as well as between $\ell_{011}$ and $\ell_{100}$. Moreover, around the origin, we draw the edges
\begin{equation}
\{01, 10\}, \{011, 100\}, \{11, 20\}, \{111, 200\}, \{21, 00\}, \{211, 000\}.
\end{equation}
This embedding is not a Freudenthal embedding, since two nonequivalent rays (e.g. $011\cdots$ and $1000\cdots$) may converge to the same point. However, we can easily modify this example to obtain a Freudenthal embedding of the same graph, by replacing the binary expansion $2^{-i}\ell_i$ in (A.8) by ternary expansion $2 \cdot 3^{-i}\ell_i$, as drawn in the top right of Figure 6. In this latter example, the set of ends is mapped onto a Cantor subset of the unit circle. There is now a face $f$ of $(G, \phi)$ that contains the exterior of the unit disc but also extends inside the unit disc; its boundary consists of a countably infinite collection of disjoint double rays together with all of $\phi(G) \setminus \phi(G)$. Note that here $\partial f \cap \phi(G)$ is dense in $\partial f$.

4. The graphs in the bottom row of Figure 6 are examples of Freudenthal embeddings of graphs with more than one end, in which each face is bounded by a cycle.
The left graph has two ends (and can easily be modified to produce any finite or countable number of ends), while the right graph has uncountably many ends.

Let us now say that faces $f_1$ and $f_2$ border each other in an edge $e$ if $e$ lies in the boundary of both faces, with $f_1$ on one side of $e$ and $f_2$ on the other. Since $G$ is 2-connected, it is not hard to see that no face borders itself; rather, each edge $e$ lies in the common boundary of precisely two faces, which border each other in $e$. If $G$ is moreover 3-connected, then two faces border each other in at most one edge. So let us assume henceforth that $G$ is 3-connected. For each edge $e$ of $G$, we let $e^* := \{f_1, f_2\}$ denote the pair of faces of $G$ that border each other in $e$. We then define the geometric dual $G^* = (V^*, E^*)$ of the plane graph $(G, \phi)$ to be the graph whose vertex set $V^*$ is the set of all faces of $(G, \phi)$ that have at least one edge in their boundary [i.e., fall in case (i) or (ii) of Theorem A.9] and whose edge set is $E^* := \{e^*: e \in E\}$. Note that the edge sets $E$ and $E^*$ are in bijection under $e \mapsto e^*$; for any subset $B \subseteq E$ we denote by $B^* = \{e^*: e \in B\}$ the corresponding subset of $E^*$.

We can now prove that geometric duals, in this generality, are always abstract duals:

**Theorem A.10. (Geometric duals are abstract duals)** Let $G$ be a locally finite 3-connected graph, and let $\phi: G \to S$ be a self-accumulation-free pointed embedding of $G$ in the sphere. Let $G^*$ be the geometric dual associated with this embedding. Then:

(i) $G^*$ is connected.

(ii) $G^*$ is an abstract dual of $G$.

(iii) $G^*$ is locally finite if and only if every face of $G$ that has at least one edge in its boundary is bounded by a cycle.

**Proof.** For each vertex $v$ of $G$, let $E_v$ denote the set of edges incident to $v$. Since $E_v$ is finite (say, of cardinality $n$) and $\phi$ is self-accumulation-free, basic topological considerations (based on the Jordan curve theorem) imply that $\phi(v)$ is surrounded by exactly $n$ faces that border each other pairwise (in cyclic order) in the edges of $E_v$, and that $E_v^*$ is a cycle in $G^* = (V^*, E^*)$.

(i) Consider any pair $v^*, w^* \in V^*$. By definition, $v^*$ and $w^*$ are faces of $(G, \phi)$ that contain at least one edge in their boundary; so let $v$ and $w$, respectively, be any endvertex of any such edge. Since $G$ is connected, there exists a path $v = v_1, v_2, \ldots, v_k = w$ in $G$. Since each cycle $E_{v_i}^*$ is connected, and $E_{v_i}^* \cap E_{v_{i+1}}^* = \{\{v_i, v_{i+1}\}\} \neq \emptyset$, the set $\bigcup_{i=1}^k E_{v_i}^*$ is connected; moreover, $v^*$ belongs to $E_{v_1}^*$ and $w^*$ belongs to $E_{v_k}^*$. So there exists a path in $\bigcup_{i=1}^k E_{v_i}^*$ from $v^*$ to $w^*$.

(ii) Consider any cycle $C$ in $G$. Then $\phi(C)$ is a circle in $\phi(G)$ that (by the Jordan curve theorem) partitions $V^*$ into two sets (call them $V_1^*$ and $V_2^*$) corresponding

---

38 In fact, it is easy to see that 2-edge-connectedness and 3-edge-connectedness are sufficient, respectively, for these statements.
to faces lying in the two components of $S \setminus \phi(C)$. Since $C$ consists exactly of those edges $e \in E$ that lie on the boundary of some pair of faces $v_1^* \in V_1^*$ and $v_2^* \in V_2^*$, it follows that $V_1^*$ and $V_2^*$ are both nonempty, so that the set $C^*$ is a cutset of $G^*$.

For each $v \in C$, the set $E_v \cap C$ consists of precisely two edges (call them $e$ and $f$) such that $e$ borders $v_1^* \in V_1^*$ and $v_2^* \in V_2^*$ and $f$ borders $w_1^* \in V_1^*$ and $w_2^* \in V_2^*$. Then $E_v^* \setminus C^*$ is the disjoint union of two paths $P_1^*, P_2^*$ in $G^*$ that link $v_1^*$ and $w_1^*$ in $V_1^*$, and $v_2^*$ and $w_2^*$ in $V_2^*$, respectively. Doing this for all $v \in C$, it follows that the set of all vertices in $V_1^*$ (resp. $V_2^*$) that are incident to an edge of the cutset $C^*$ is connected in $V_1^*$ (resp. $V_2^*$). Therefore $C^*$ is a minimal cutset of $G^*$.

Conversely, if $B$ is a finite subset of $E$ such that $B^*$ is a cutset of $G^*$, then for each $v \in V$, the cutset $B^*$ must intersect $E_v^*$ (which is a cycle in $G^*$) an even number of times (that is, $B^* \cap E_v^*$ has even cardinality). It follows that each vertex of $G$ is incident to an even number of edges of $B$, and hence $B$ contains a cycle $C$. But we have just proven that $C^*$ is a cutset in $G^*$. Therefore, if $B^*$ is a minimal cutset of $G^*$, we necessarily have $B = C$.

(iii) By Theorem A.9 for every face $f$ of $(G, \phi)$ that has at least one edge in its boundary (i.e., belongs to $V^*$), its boundary $\partial f$ must either be a cycle or contain a double ray. Clearly, $f$ has finite degree in $G^*$ in case $\partial f$ is a cycle, and has countably infinite degree in case $\partial f$ contains a double ray. □

In particular, whenever $G$ is a locally finite 3-connected graph and $\phi: G \to S$ is a self-accumulation-free pointed embedding such that every face of $G$ is bounded by a cycle, the geometric dual $G^*$ is a locally finite abstract dual of $G$, so that all the nice properties stated in Theorem A.3 ensue.

Conversely, let us now show that if a locally finite 3-connected graph $G$ has a locally finite abstract dual, then $G$ has a unique self-accumulation-free pointed embedding, and the geometric dual associated with this embedding coincides with the abstract dual:

**Theorem A.11. (Embeddings of graphs with locally finite abstract duals)**

Let $G = (V, E)$ be a locally finite 3-connected graph that has a locally finite abstract dual $G^\dagger = (V^\dagger, E^\dagger)$. Then:

(i) $G$ is planar.

(ii) In the Freudenthal embedding of $G$, each face is bounded by a cycle, and the geometric dual $G^*$ of this embedding coincides with the (unique) abstract dual $G^\dagger$.

(iii) The Freudenthal embedding is (up to homeomorphism) the only self-accumulation-free pointed embedding of $G$ in the sphere.

**Proof.** (i) is proven in [68, Theorem 9.3]. As a consequence of (i), Theorem A.8 guarantees the existence and uniqueness (up to homeomorphism) of the Freudenthal embedding.

As a preliminary to (ii) and (iii), consider any self-accumulation-free pointed embedding of $G$ in the sphere, and let $G^*$ be the geometric dual associated with this
embedding. By Theorem A.10, $G^*$ is an abstract dual of $G$. But by Theorem A.3 (iii), $G$ has a unique abstract dual $G^\dagger$. Hence $G^\dagger = G^*$.

(ii) For the Freudenthal embedding, Theorem A.9 guarantees that every face has at least one edge in its boundary. And since $G^\dagger = G^*$ is locally finite, Theorem A.10 (iii) implies that every face is bounded by a cycle.

(iii) Let $R_1, R_2$ be inequivalent rays in $G$. Then there exists a finite set of edges $B$ in $G$ such that the tails of $R_1$ and $R_2$ lie in different connected components of $G \setminus B$. Making $B$ smaller if necessary, we may assume without loss of generality that $B$ is a minimal set with this property. It is not hard to see that $B^\dagger$ is a cycle in $G^\dagger$.

Since $G^\dagger = G^*$, $B^\dagger$ is a cycle in $G^\dagger$. By (ii), each of the vertices of the cycle $B^\dagger$ is a face of $(G, \phi)$ whose boundary is a cycle in $G$. The sum modulo 2 of all these cycles forms a pair of disjoint cycles in $G$, each of which separates the tail of $\phi(R_1)$ from the tail of $\phi(R_2)$. As a result, these tails necessarily converge to different endpoints. Since this holds for any pair of inequivalent rays, we conclude that $G^\dagger$ must be the Freudenthal compactification, and $\phi$ the Freudenthal embedding. □

Until now we have defined dual graphs — whether in the abstract or geometric sense — simply as abstract graphs without a given embedding in the plane. However, given an embedding of a planar graph, there is a natural way to embed its dual. Let $G$ be a locally finite 3-connected graph, let $\phi: G \to S$ be a self-accumulation-free pointed embedding of $G$ in which each face is bounded by a cycle, and let $G^*$ be the geometric dual of $(G, \phi)$. If $e$ is an edge of $G$, we denote by $\hat{e}$ the interior of the edge $e$ in the topological realization of $G$ (namely, the edge without its endvertices). We then say that an embedding $\phi^*$ of $G^*$ is a dual embedding to $(G, \phi)$ if

(i) For each $v^* \in V^*$, we have $\phi^*(v^*) \in v^*$.
(ii) For each $e \in E$, the arc $\phi^*(e^*)$ intersects $\phi(G)$ in a single point, which lies on $\phi(\hat{e})$.

Less formally, (i) says that each vertex of $G^*$ is represented by a point lying in the corresponding face of $(G, \phi)$, and (ii) says that two such points that lie in faces that border each other in an edge $e$ are linked by a dual edge $e^*$ that crosses $e$ in a single interior point and is otherwise disjoint from $\phi(G)$. We then have:

**Proposition A.12. (Dual embeddings)** Let $G$ be a locally finite 3-connected graph, and let $\phi: G \to S$ be a self-accumulation-free pointed embedding of $G$ such that each face of $(G, \phi)$ is bounded by a cycle. Let $G^*$ be the geometric dual associated with this embedding. Then there exists a dual embedding $\phi^*$ of $G^*$. Moreover, for each such $\phi^*$:

(i) $\phi^*$ is self-accumulation-free and pointed, and each face of $(G^*, \phi^*)$ is bounded by a cycle.
(ii) $G$ is a geometric dual of $(G^*, \phi^*)$ and $\phi$ is a dual embedding of $G$.
(iii) The sets of endpoints $\phi(G) \setminus \phi(G)$ and $\phi^*(G^*) \setminus \phi^*(G^*)$ coincide.
If \((G, \phi)\) and \((G^*, \phi^*)\) are as in Proposition \ref{prop:dual}, then we say that they form a geometric dual pair.

**Proof of Proposition \ref{prop:dual}**. To prove the existence of a dual embedding \(\phi^*\) of \(G^*\), we begin by choosing a point \(\phi^*(v^*) \in v^*\) for each face \(v^*\) of \((G, \phi)\), and a point \(x\) on \(\phi(\hat{e})\) for each edge \(e \in E(G)\). Let \(x_1, \ldots, x_n\) be the points chosen on the edges of the cycle that bounds \(v^*\). It follows from basic topological considerations\(^{39}\) that we can connect \(\phi^*(v^*)\) to each of the points \(x_1, \ldots, x_n\) by continuous arcs that are disjoint except at their common endpoint \(\phi^*(v^*)\) and that lie entirely in \(v^*\) except for their endpoints \(x_1, \ldots, x_n\). Now if \(e \in E(G)\) lies on the boundary of faces \(v_1^*\) and \(v_2^*\) and \(x\) is the chosen point on \(\phi(\hat{e})\), then the concatenation of the arcs from \(\phi^*(v_1^*)\) to \(x\) and from \(x\) to \(\phi^*(v_2^*)\) is an arc from \(\phi^*(v_1^*)\) to \(\phi^*(v_2^*)\) that we take as our definition of \(\phi^*(e^*)\).

To see that \(\phi^*\) is self-accumulation-free, it suffices to show that for each point \(x \in G^*\) we can remove a finite set of edges \(E_0^*\) from \(G^*\) such that \(\phi^*(x) \notin \phi^*(G^*\setminus E_0^*)\). If \(x = v^*\) is a vertex of \(G^*\), then it suffices to remove the set \(E_0^*\) of all edges incident to \(v^*\). Then \(\phi^*(G^*\setminus E_0^*)\) is contained in the complement of the face \(v^*\) and hence its closure cannot contain \(\phi^*(x)\). If \(x\) lies on an edge \(e^*\) connecting two vertices \(v_1^*\) and \(v_2^*\) of \(G^*\), then we remove the set \(E_0^*\) of all edges incident to \(v_1^*\) or \(v_2^*\). Then \(\phi^*(G^*\setminus E_0^*)\) is contained in the complement of the open set formed by the union of the faces \(v_1^*\) and \(v_2^*\) and \(\phi(\hat{e})\), while \(\phi^*(x)\) lies inside this open set.

To see that \(\phi^*\) is pointed, let \(R^*\) be a ray in \(G^*\) consisting of consecutive vertices \(v_0^*, v_1^*, \ldots\) connected by edges \(e_1^*, e_2^*, \ldots\). Then \(v_0^*, v_1^*, \ldots\) are faces of \((G, \phi)\) which by assumption are bounded by cycles \(C_0, C_1, \ldots\) such that \(C_{k-1} \cap C_k = e_k\). Adding these cycles modulo 2 yields a double ray consisting of two rays \(R_1, R_2\) in \(\phi(G)\). Moreover, we can construct a third ray \(R_3\) that lies in the union of the cycles \(C_0, C_1, \ldots\), passes through each of the edges \(e_1, e_2, \ldots\), and has an infinite intersection with both \(R_1\) and \(R_2\). In particular, the rays \(R_1, R_2\) and \(R_3\) are all equivalent and thus, as \(\phi\) is pointed, they all converge to the same point \(x \in S\). It follows that for each open ball around \(x\), there exists an \(n\) such that the sum modulo 2 of the cycles \(C_n, C_{n+1}, \ldots\) lies in this ball. But this is a double ray which together with the point \(x\) forms a circle in \(S\) that contains the tail of the ray \(R^*\) in its interior. Any open ball around \(x\) thus eventually contains the tail of \(R^*\), which proves that \(R^*\) converges to \(x\).

This proves not only that \((G^*, \phi^*)\) is pointed, but also that every point in \(S\) that is the limit of some ray in \((G^*, \phi^*)\) is also the limit of some ray in \((G, \phi)\). It follows that the set of endpoints of \((G^*, \phi^*)\) is contained in the set of endpoints of \((G, \phi)\): that is, \(\phi^*(G^*) \setminus \phi^*(G^*) \subseteq \phi(G) \setminus \phi(G)\).

\(^{39}\) Indeed, this can be proven by repeated application of the Jordan curve theorem and the following basic topological fact: Let \(C \subset S\) be a circle, which by the Jordan curve theorem divides \(S\) into two connected open sets \(U, V\); let \(x, z\) be two distinct points on \(C\), and let \(y \in U\). Then there exists a continuous arc from \(x\) to \(z\) that lies, except for its endpoints, entirely in \(U\) and passes through \(y\).
To complete the proof, it now suffices to show that each face of \((G^*, \phi^*)\) is bounded by a cycle [completing the proof of (i)] and contains a single vertex of \((G, \phi)\) [which implies (ii)]. Then (ii) together with what we already got implies the reverse inclusion \(\overline{\phi(G)} \setminus \phi(G) \subseteq \overline{\phi^*(G^*)} \setminus \phi^*(G^*)\), so we conclude that (iii) holds.

Since every face of \((G, \phi)\) is bounded by a cycle, the set \(\phi(G) \setminus \phi^*(G^*)\) does not separate \(S\), so the same is true for \(\overline{\phi^*(G^*)} \setminus \phi^*(G^*)\) which is contained in it. It follows that each face \(f\) of \((G^*, \phi^*)\) has at least one edge in its boundary and hence by Theorem A.9 is either a cycle \(C^*\) or contains a double ray \(R^*\). For each vertex \(v^*\) in \(C^*\) or \(R^*\) we have that \(\phi^*(v^*)\) lies inside a cycle \(C\) of \(G\) that crosses two consecutive edges of \(C^*\) or \(R^*\). This cycle \(C\) can have only one vertex in \(f\) since otherwise there would be an edge of \((G, \phi)\) that lies entirely in \(f\), which contradicts our construction of \(\phi^*\). It follows that all edges in \((G, \phi)\) that cross an edge in \(C^*\) or \(R^*\) are incident to one and the same vertex \(v\) of \(G\) with \(\phi(v) \in f\). Since \(G\) is locally finite, we can rule out the double ray so we conclude that \(f\) is bounded by a cycle \(C^*\) and contains a single vertex of \((G, \phi)\). \(\square\)

Now recall the definition of vertex and edge accumulation points (just before Lemma A.7). Following [68, 69], let us say that an embedding \(\phi: G \to \mathbb{R}^2 \cup \{\infty\} \cong \mathbb{S}\) is VAP-free (resp. EAP-free) if \(\phi(G) \subset \mathbb{R}^2\) and \(\phi\) has no VAPs (resp. EAPs) in the finite plane \(\mathbb{R}^2\). Note that if \(G\) has at most finitely many isolated vertices (in particular, if \(G\) is connected), then every VAP is also an EAP. For such graphs, EAP-free embeddings are automatically self-accumulation-free and pointed.

An EAP-free embedding \(\phi\) of an infinite, connected, locally finite graph \(G\) can be extended to an embedding \(\overline{\phi}\) of the one-point compactification \(G^*\) of \(G\) by setting \(\overline{\phi}(\infty) = \infty\). Conversely, if the one-point compactification \(G^*\) of \(G\) can be embedded in the sphere \(\mathbb{S} \cong \mathbb{R}^2 \cup \{\infty\}\), then without loss of generality we may assume \(\overline{\phi}(\infty) = \infty\), yielding an EAP-free embedding of \(G\). We then have:

**Proposition A.13. (Graphs with one end)** Let \(G\) be a locally finite 3-connected planar graph.

(i) If \(G\) has at most one end, then it has an EAP-free embedding, which is unique up to homeomorphism and coincides with the Freudenthal embedding.

If in addition \(G\) has a locally finite abstract dual \(G^\dagger\), then:

(ii) \(G\) has an EAP-free embedding if and only if it has at most one end; and in this case \(G\) and \(G^\dagger\) can be represented as a geometric dual pair \((G, \phi), (G^*, \phi^*)\) such that both \((G, \phi)\) and \((G^*, \phi^*)\) are EAP-free.

---

40 Since \(\phi(G) \subset \mathbb{R}^2\) and \(\phi\) has no EAPs in \(\mathbb{R}^2\), it must be self-accumulation-free. Since every ray has at least one accumulation point in \(\mathbb{R}^2 \cup \{\infty\}\) and no accumulation points in \(\mathbb{R}^2\), all rays converge to \(\infty\). It seems to us that the concepts “self-accumulation-free” and “pointed” can now largely replace the more restricted notion of being EAP-free.
Proof. (i) If $G$ is finite, then it obviously has an EAP-free embedding. If $G$ is infinite with one end, then it has an EAP-free embedding since its one-point compactification coincides with its Freudenthal compactification, which by Theorem A.8(a) can be embedded in the sphere. And by Theorem A.8(b) this embedding is unique up to homeomorphism.

(ii) If $G$ has a locally finite abstract dual $G^\dagger$, then Theorem A.11(iii) and Lemma A.6 tell us that the Freudenthal compactification is the only pointed compactification of $G$ that is embeddable in the sphere; therefore, such a graph has an EAP-free embedding if and only if it has at most one end. (This statement can also be found in [69, Theorem 5.9].) By Proposition A.12, $G$ and $G^\dagger$ can be represented as a geometric dual pair $(G, \phi)$, $(G^*, \phi^*)$; and Proposition A.12(iii) then implies in particular that if $(G, \phi)$ is EAP-free, then so is $(G^*, \phi^*)$. □

A.4. Special embeddings. In this section we discuss embeddings with special “nice” properties: notably, straight-line embeddings and periodic embeddings.

We begin by citing a result of Thomassen [68] on straight-line embeddings with convex faces:

Proposition A.14. (Convex embeddings) Let $G$ be a locally finite, 3-connected graph with one end. Then there exists an EAP-free embedding of $G$ in which every edge is a straight line segment and every face is convex. If in addition $G$ has a locally finite abstract dual $G^\dagger$, then there exists an EAP-free embedding of $G$ in which every face is a convex polygon.

Proof. By Proposition A.13(i), $G$ has an EAP-free embedding in the plane. Then [68, Theorems 7.4 and 8.6] imply that $G$ has an EAP-free embedding in which every edge is a straight line segment and every face is convex. If in addition $G$ has a locally finite abstract dual $G^\dagger$, then Theorem A.11 guarantees that every face is bounded by a cycle. □

Using Proposition A.14, we can deduce that $G$ and $G^\dagger$ can be jointly embedded as a geometric dual pair such that edges are represented by straight-line segments in both graphs:

Proposition A.15. (Straight-line embedding of dual pair) Let $G$ be a locally finite, 3-connected graph with one end. Assume that $G$ has a locally finite abstract dual $G^\dagger$. Then there exist EAP-free embeddings $\phi$ and $\phi^\dagger$ of $G$ and $G^\dagger$ in the (Euclidean) plane such that $(G, \phi)$ and $(G^\dagger, \phi^\dagger)$ form a geometric dual pair and each edge of $(G, \phi)$ and $(G^\dagger, \phi^\dagger)$ is a straight line segment.

Proof. By Theorem A.11 and Proposition A.12, we may embed $G$ and $G^\dagger$ as a geometric dual pair. Now we can also draw a graph $H$ in the plane such that the vertex set of $H$ is the union of the vertex sets of $G$ and $G^\dagger$ and two vertices $v \in G$ and $w \in G^\dagger$ are joined by an edge in $H$ when $v$ and $w$ are the endpoints of an
edge in $G$ and its dual edge in $G^\dagger$; thus, $H$ is a quadrangulation, with $G$ and $G^\dagger$ being its sublattices whose edges connect opposing vertices of the quadrilaterals of $H$. Then obviously $H$ is planar, has an EAP-free embedding in the plane, and has a locally finite geometric (and hence abstract) dual. Since two faces of $H$ never have the property that two non-adjacent vertices of $H$ both lie on the boundary of both faces, we see that $H$ is 3-connected. By Proposition A.13(ii), $H$ has one end, so applying Proposition A.14 to $H$ we see that $H$ has an EAP-free embedding such that each face is a convex polygon (with four corners). Connecting opposite corners of these faces by straight line segments, we obtain the required straight-line embeddings of $(G,\phi)$ and $(G^\dagger,\phi^\dagger)$ as a geometric dual pair. □

Remark. We do not know whether $G$ and $G^\dagger$ can be jointly embedded as a geometric dual pair such that edges are represented by straight line segments in both graphs and faces are convex in both graphs.

We next turn our attention to embeddings of quasi-transitive graphs. Our main graphs of interest — namely, the graph $G$ and its sublattices $G_0$ and $G_1$ arising in Theorem A.17 — are all locally finite, 3-connected, quasi-transitive, planar graphs with one end that have a locally finite abstract dual $G^\dagger$. Remarkably, this latter condition turns out to be superfluous:

**Proposition A.16. (Quasi-transitive graphs with one end)** Let $G$ be a locally finite, 3-connected, quasi-transitive, planar graph with one end. Then:

(i) $G$ has a locally finite abstract dual $G^\dagger$.

(ii) The Freudenthal compactification is the only pointed compactification of $G$.

(iii) The Freudenthal embedding of $G$ is EAP-free (when the endpoint is taken to map to $\infty$), and each face is bounded by a cycle.

**Proof.** (ii) is trivial because $G$ has one end (and is 2-connected). To prove (iii) and (i), consider the Freudenthal embedding of $G$. By Proposition A.13(i), it is EAP-free. By Theorem A.9, no face is bounded by endpoints alone; and since $G$ has one end, a face boundary containing infinitely many edges must consist of a single double ray together with the one endpoint. But quasi-transitivity then implies (see [6, Theorem 2.3] or [49, Theorem 8(1)]) that no such infinite face can exist, i.e., every face is bounded by a cycle. From Theorem A.10 we then conclude that $G$ has a locally finite abstract dual $G^\dagger$. □

Quasi-transitivity is essential here, as the example of $\mathbb{N} \times \mathbb{Z}$ (see Example 2 after Theorem A.9) shows. The assumption that $G$ has one end is also essential: let $H$ be the graph with vertex set $\mathbb{Z}$ and edges $\{n, n+1\}$ and $\{2n, 2n+2\}$ for all $n \in \mathbb{Z}$, let $G$ be the ladder graph $\{0,1\} \times H$, and let $\phi$ be the obvious embedding. Then $G$ is quasi-transitive and 3-connected and has two ends; $\phi$ is EAP-free and maps both ends to the point $\infty$ (i.e., this is a non-Freudenthal embedding); and $(G,\phi)$ has a pair of faces that are each bounded by a double ray together with the point $\infty$. 
When planar graphs are quasi-transitive, it is natural to ask if they can be embedded in a periodic way in the plane. This is not true if one restricts oneself to the Euclidean plane, but remarkably, it turns out to be correct if one also considers the hyperbolic plane. The following result has been proven in [3, Theorem 4.2] (see also [67, Theorem 1]):

**Theorem A.17. (Periodic embeddings)** Every locally finite, 3-connected, quasi-transitive planar graph $G$ with one end can be embedded in the Euclidean plane $\mathbb{R}^2$ or hyperbolic plane $\mathbb{H}^2$ such that every automorphism of $G$ corresponds to an isometry of $\mathbb{R}^2$ or $\mathbb{H}^2$, respectively.

We remark that in Theorem [A.17], we do not know if the embedding can be chosen in such a way that, moreover, edges are represented by straight line segments in $\mathbb{R}^2$ or $\mathbb{H}^2$. Note also that Propositions [A.14] and [A.15] talk about embeddings such that edges are straight line segments in the Euclidean geometry. We are not aware of results about straight-line embeddings in the hyperbolic geometry.

**A.5. Some examples.** Finally, let us describe a method for creating examples of graphs satisfying the assumptions of Theorem [1.1]—i.e., locally finite, 3-connected, quasi-transitive triangulations with one end—and their duals. All our examples come naturally with a periodic straight-line embedding in the Euclidean or hyperbolic plane.

Let $p, q \geq 3$ be integers and let $ABC$ be a triangle whose angles (in anticlockwise order) at the corners $A, B, C$ are $\pi/p$, $\pi/q$, and $\pi/2$, respectively. Such a triangle can be constructed in either the sphere, the Euclidean plane, or the hyperbolic plane, depending on whether $1/p + 1/q + 1/2$ is larger than, equal to, or less than 1, respectively. By reflecting the triangle $ABC$ in one of its edges and continuing this process, we can cover the whole space alternately by copies of $ABC$ and its mirror image [12, Section 2]. This yields a planar graph with vertices of types $A$, $B$ and $C$ that are of degree $2p$, $2q$ and 4, respectively. In particular, each vertex of type $C$ is adjacent to two vertices of types $A$ and $B$ each, in alternating order. We may view the $A$ and $B$ sublattices as planar graphs in their own right by erasing the vertices of type $C$ and viewing the four edges emanating from $C$ as two straight edges crossing each other in $C$, where one connects two $A$’s and the other connects two $B$’s. This yields two regular tessellations that are geometric duals of each other. In the tessellation formed by the $A$ vertices, each vertex has degree $p$ and each face is a regular polygon with $q$ edges. This regular tessellation is denoted by the Schlafli symbol $\{q, p\}$ [12]. Likewise, the dual $B$ lattice has the Schlafli symbol $\{p, q\}$.

In particular, the tessellations with Schlafli symbol $\{3, p\}$ (with $p \geq 3$) are regular triangulations of the sphere, the Euclidean plane, or the hyperbolic plane, depending on whether $p$ is less than, equal to, or larger than 6, respectively. It is easy to see that $\{3, p\}$, as a graph, is 3-connected and vertex-transitive. It is finite for $p \leq 5$ and infinite for $p \geq 6$. In particular, Theorem [1.1] applies when $G_0 = \{3, p\}$ with $p \geq 6$. The case $p = 6$, which is the only Euclidean tessellation in this class, yields
\(G_0 = \text{triangular lattice, } G_1 = \text{hexagonal lattice, and } G = \text{diced lattice. The cases } p > 6 \text{ yield hyperbolic tessellations. The graphs } \{3, 6\} \text{ and } \{3, 7\} \text{ and their duals are drawn in Figure 2(b,d).}

The (dual) tessellations with Schl"afli symbol \(\{p, 3\}\) are planar Cayley graphs in which every vertex has degree three. A full classification of graphs with these properties can be found in [29]. In particular, [29, Table 1, items 12–19] lists those that are 3-connected and have at most one end. Note that all these graphs are vertex-transitive.

More general examples of quasi-transitive triangulations satisfying the assumptions of Theorem 1.1 can be constructed by starting with any regular tessellation and dividing the basic polygon into triangles in some suitable way, so that the resulting graph is 3-connected. It would go too far to attempt here a full classification of the class of tessellations covered by Theorem 1.1.

Acknowledgements

We wish to thank Youjin Deng, Jesper Jacobsen, and Jesús Salas for many helpful conversations over the course of this work. We also thank Agelos Georgakopoulos, Sebastian M"uller, Bruce Richter, Jan Seidler, and Carsten Thomassen for helpful discussions and correspondence concerning infinite graphs and planar topology; Neal Madras and Gordon Slade for correspondence concerning self-avoiding walks; and Youjin Deng, Kun Chen and Yuan Huang for sharing with us their preliminary Monte Carlo data.

The research of R.K. and J.M.S. was supported in part by the grants GAČR 201-09-1931 and P201/12/2613. The research of A.D.S. was supported in part by U.S. National Science Foundation grant PHY–0424082.

References


