The BFACF algorithm and knotted polygons

1991 J. Phys. A: Math. Gen. 24 5553

(http://iopscience.iop.org/0305-4470/24/23/021)

J. Phys. A: Math. Gen. **24** (1991) 5553–5567. Printed in the UK

The BFACF algorithm and knotted polygons

E J Janse van Rensburg† and S G Whittington‡

† Supercomputer Computations Research Institute, Florida State University, Tallahassee, FL 32306-4052, US

‡ Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 1A1, Canada

Received 11 March 1991

Abstract. The BFACF algorithm applied to polygons involves sampling on a Markov chain whose state space is the set of all polygons. In three dimensions, for the simple cubic lattice, we prove that the ergodic classes of this Markov chain are the knot classes of the polygons.

1. Introduction

Monte Carlo treatments of the self-avoiding walk and related problems fall into two broad classes. One of these could be described as *walk growing methods* in which the walk is constructed step by step. For a discussion of this approach see e.g. Hammersley and Handscomb (1964), chapter 10. The other general approach is to make changes in the walk and produce a correlated sequence of walks. This approach is often called *Metropolis style* Monte Carlo (Metropolis *et al* 1953) and involves sampling along a realization of a Markov chain, whose (unique) limit distribution is the desired distribution. For example, if the problem being studied is the pure n-step self-avoiding walk, the state space of the Markov chain should be the set of all n-step self-avoiding walks and the limit distribution should be uniform. The limit distribution must be unique and not depend on the initial state of the realization of the Markov chain (Hammersley and Handscomb 1964).

Early attempts in this direction include papers by Verdier and Stockmayer (1962), Kron (1965), Bluestone and Vold (1965), Monnerie and Geny (1969) and Lal (1969), although some of these workers had in mind a simulation of the dynamics of the polymer being modelled by the walk. There has been a recent renewal of interest in this kind of approach, both from a theoretical and from a practical point of view. Madras and Sokal (1987) proved that no such method which is length conserving (i.e. for fixed n) and uses only local moves can be ergodic, which implies that the limit distribution will depend on the initial state of the realization.

This paper is concerned with the BFACF algorithm (Berg and Foester 1981, Aragao de Carvalho *et al* 1983, Aragao de Carvalho and Caracciolo 1983) which simulates walks with variable length and *fixed endpoints* in the hypercubic lattice. When applied to self-avoiding walks with fixed endpoints in the square lattice the underlying Markov chain is known to be ergodic (Madras 1986). The BFACF algorithm is not related to the Berretti–Sokal algorithm (Berretti and Sokal 1985) which simulates walks with free endpoints and variable length. The dynamical behaviour of the BFACF algorithm is known to be considerably worse than the Berretti–Sokal algorithm (for recent progress in understanding the dynamical behaviour of these algorithms see, for example, Caracciolo and Sokal 1986, Sokal and Thomas 1988, Caracciolo *et al* 1990a, b, 1991).

The algorithm can be applied to polygons but in this case ergodicity questions are rather more tricky. For the square lattice Madras (1986) has shown that the algorithm is ergodic but, in three dimensions, it is easy to see that the BFACF moves will not convert a knotted polygon to an unknotted polygon. This is enough to show that the algorithm is not ergodic. In this paper we prove that the ergodic classes are the knot types. That is, if two polygons are of the same knot type they can be interconverted by BFACF moves, but not otherwise. This result is interesting from two points of view. It means that the BFACF algorithm cannot be applied (as it stands) to sample uniformly from the space of all polygons but, perhaps more importantly, it is a convenient algorithm for investigating the properties of polygons of fixed knot type. Elsewhere we make use of this feature to investigate the dimensions of polygons of fixed knot type (Janse van Rensburg and Whittington 1991).

2. Definitions

2.1. The BFACF Algorithm

An unrooted self-avoiding polygon, or *polygon* ω, in any lattice, is a sequence of lattice sites $\omega_0, \omega_1, \omega_2, \ldots, \omega_n$, and associated edges (ω_i, ω_{i+1}) such that: $\omega_0 = \omega_n$, and ω_i and ω_{i+1} are nearest neighbours in the lattice, and $\omega_1, \omega_2, \ldots, \omega_n$ are all distinct. If $\omega_0 \neq \omega_n$ then we call ω a *walk*. Let Z^d be the d-dimensional hypercubic lattice. Let $\{e_i\}_{i=1}^{d}$ be the set of orthogonal unit vectors in Z^d. The BFACF algorithm is a local stochastic process which operates on paths (any sequence of edges) in the hypercubic lattice. It generates statistical ensembles of paths with a Boltzmann distribution (Berg and Foester 1981). The algorithm was first applied to 'bosonic' walks (Brownian walks) and 'fermionic' walks (Brownian walks without 'spikes') by Berg and Foester (1981), before it was applied to the self-avoiding walk by Aragao de Carvalho *et al* (1983) and Aragao de Carvalho and Caracciolo (1983).

Let P be the space of equivalence classes of polygons modulo a translation in Z^d, and let W_{xy} be the set of all walks connecting the lattice sites x and y in Z^d. (In the subsequent discussions, we refer to elements of P as 'polygons'; since these are in fact equivalence classes modulo a translation, we see that they are *unrooted polygons*. Observe that these are not the same as walks which start and terminate at neighbouring lattice sites, which correspond to *rooted polygons*.) The BFACF algorithm operates on ω (which is either a walk in W_{xy} or an unrooted polygon in P) by attempting one of the local deformations shown in figure 1, at a randomly selected location on edges *which are present* in ω. If successful, these elementary operations result in a length change in ω which is either -2, 0 or $+2$. Observe that if ω is a walk, then the local deformations cannot move its endpoints (which are therefore fixed in the lattice). In contrast, an unrooted polygon can be translated anywhere, as we shall prove later in this paper (proposition 3.10). For a detailed description of the algorithm, see for example Caracciolo *et al* 1990a (which contains in addition to the usual implementation also a Metropolis style implementation of the algorithm).

The algorithm involves sampling along a realization of a Markov chain, the state space of which is the set of all polygons or walks with fixed endpoints. In two a)

b)

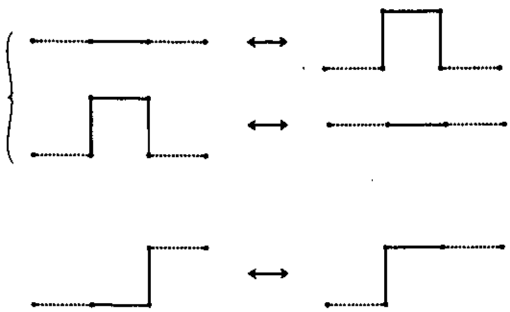

Figure 1. The elementary moves of the BFACF algorithm.

dimensions the Markov chain is known to be irreducible (Madras 1986) and in three dimensions it is known to be reducible (Madras and Sokal 1987). Let the ergodicity classes in three dimensions be $\mathcal{E}_i, i = 1, 2, \ldots$. If the algorithm is applied to polygons then it is known that the limit distribution of the Markov chain is not unique and depends on the initial state of the realization. In general, we can write the limit distribution as (Berg and Foester 1981)

$$\pi_i(\omega) = Z_i(\beta)^{-1}|\omega|\beta^{|\omega|} \tag{2.1}$$

where $\omega \in \mathcal{E}_i$ for some i, $|\omega|$ is the number of edges in ω, β is an adjustable parameter and the normalization factor $Z_i(\beta)$ is given by

$$Z_i(\beta) = \sum_{\omega \in \mathcal{E}_i} |\omega|\beta^{|\omega|}. \tag{2.2}$$

The set $\mathcal{E}_i$ is determined by the initial polygon or walk in the realization.

The transition probability matrix $\boldsymbol{P} = \{p(\omega \to \nu)\} = \{p_{\omega\nu}\}$ has the following properties in its ergodic classes: (1) For each $\omega, \nu \in \mathcal{E}_i$ there exists an $m \geqslant 0$ such that the m-step transition probability from ω to ν is positive. (2) For each polygon $\nu \in \mathcal{E}_i$, $\sum_{\omega \in \mathcal{E}_i} \pi_i(\omega)p_{\omega\nu} = \pi_i(\nu)$. It is easy to show that $\pi_i(\omega)$ is the unique limit distribution of the Markov chain with state space $\mathcal{E}_i$ and transition probability matrix $\boldsymbol{P}$ (Kemeny and Snell 1976).

In two dimensions there is only one ergodic class. In three and higher dimensions we have to determine the ergodic classes $\mathcal{E}_i$. In section 3 we consider the case in three dimensions, which is particularly interesting. In four and higher dimensions little is known about the behaviour of this algorithm.

2.2. *Knots and polygons*

Let S^1 be the circle, and consider the map $f : S^1 \to \mathcal{R}^3$, an *embedding* of a circle into Euclidean 3-space. (That is, f is one-to-one and is a homeomorphism onto its image.) We write this map as $(f; S^1, \mathcal{R}^3)$. A polygon $\omega \in \mathcal{Z}^3 \subset \mathcal{R}^3$ is a *piecewise linear* embedding of S^1 in $\mathcal{R}^3$. We call any embedding $(f; S^1, \mathcal{R}^3)$ a *knot*. An embedding $(f; S^1, \mathcal{R}^3)$ is oriented if we give an orientation to the circle. Two (piecewise linear) oriented embeddings $(f; S^1, \mathcal{R}^3)$ and $(g; S^1, \mathcal{R}^3)$ are *ambient isotopic* if there is an orientation-preserving isotopy $H : \mathcal{R}^3 \times I \to \mathcal{R}^3 \times I$ (I is the unit interval) with $H(y, t) = (h_t(y), t)$ such that h_0 is the identity, and $h_1 f = g$.

In other words, we can continuously deform $\mathcal{R}^3$ such that $(f; S^1, \mathcal{R}^3)$ is taken into $(g; S^1, \mathcal{R}^3)$. We note that two oriented polygons which are ambient isotopic do not need to have the same number of edges.

We call two (oriented) piecewise linear embeddings *equivalent* if they are ambient isotopic. We call the equivalence classes of (oriented) embeddings of the circle into Euclidean space *knot types*. In the case of polygons in $\mathcal{Z}^3$, we define a knot type as all the polygons which are in the same equivalence class when viewed as piecewise linear embeddings of the circle into $\mathcal{R}^3$.

A *projection* P of a circle on any plane $\mathcal{R}^2 \subset \mathcal{R}^3$ is called *regular* if (1) there are only finitely many multiple points $\{p_i | 1 \leqslant i \leqslant n\}$, (2) all multiple points are double points, and (3) no vertex of the knot is mapped onto a double point.

The regular projection of a knot does not determine the knot but, if we indicate at every double point the overpassing line and the orientation of the knot, we can reconstruct the knot from its regular projection. Such a projection (which allows the reconstruction of the original knot) is called a *knot projection* (Burde and Zieschang 1985).

Two knot projections are defined to be *equivalent* if they are connected by a finite sequence of Reidemeister moves (Reidemeister 1932), which are (local) operations on the knot projection. We illustrate the Reidemeister moves in figure 2. All these moves can be realised in the knot projection by an ambient isotopy of the knot. Therefore equivalent knot projections define equivalent knots. The converse is also true: equivalent knots have equivalent knot projections (Reidemeister 1923). Therefore, two knots are equivalent if and only if their projections are equivalent.

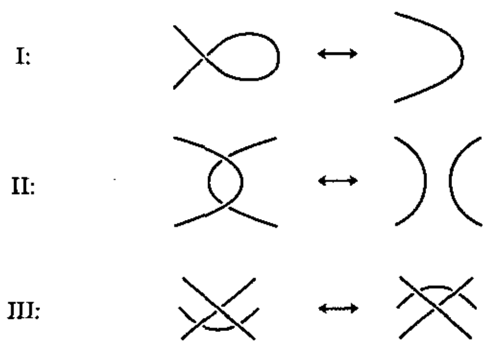

Figure 2. The three Reidemeister moves.

We now apply these ideas to polygons in the simple cubic lattice. Let $\omega \in P_n$ and $\nu \in P_m$ be two polygons in the cubic lattice where the subscript denotes the number of edges in the polygon. We must define the projections of ω and ν on a convenient reference plane in the lattice. Let $\{e_1, e_2, e_3\}$ be the set of three orthogonal unit vectors in $\mathcal{Z}^3$. Let $Q = \{xe_1 + ye_2 + ze_3 | z = 0, \quad x, y \in \mathcal{R}\}$ be the $(z = 0)$-plane in $\mathcal{R}^3$. Let $\omega \in P_n$ be any polygon in three-dimensions, and consider the projection of ω in Q. In general, ω has edges parallel to the three lattice axes, and those edges parallel to the e_3 axis will project to a single point in Q. We ignore these edges in the projection. Multiple points in the projection will occur at points with integer coordinates in Q, or will be line segments which connect points with integer coordinates in Q. The projection of a polygon in Q is in general not regular, even if we ignore edges in the e_3 direction. If we, in addition, discount vertices in the projection where two projected edges make 180° angles (that is, the only vertices in the projected polygon are those where two edges make a 90° angle), then the projection of ω in Q might be regular. If we indicate overpasses and the orientation, then we can reconstruct ω from its projection, except for the edges in ω in the e_3 direction which are lost in the projection. We call the regular projections of lattice polygons into the $(z = 0)$-plane *lattice knot projections.*

If we want to determine the ergodic classes of lattice polygons under the BFACF moves, then we must study the relation of Reidemeister moves to BFACF moves. It is obvious that, in any regular projection of ω into a plane with a normal vector with irrational direction cosines, a BFACF move corresponds at most to a finite sequence of Reidemeister moves. The ergodic classes of the algorithm are therefore subsets of the knot types of the polygon in three dimensions. If the knot types of polygons (when viewed as piecewise linear knots) are represented by K_i, then for each ergodic class in three dimensions, $\mathcal{E}_i$, there exists a j such that $\mathcal{E}_i \subset K_j$. We therefore have the following proposition:

Proposition 2.1. Let $d = 3$ and let ω and ν be two polygons in the set $P = \cup_n P_n$. If ω and ν are in the same ergodic class, then they have the same knot-type.

Proof. A BFACF move corresponds to a continuous deformation of the pair $(\omega, \mathcal{R}^3)$, if we consider the polygon to be embedded in Euclidean 3-space. □

Consider a self-avoiding walk in $\mathcal{Z}^3$ which is a member of W_{xy}. Suppose that the projection of this walk in the $(z = 0)$-plane is a self-avoiding walk. A BFACF move on the projection (that is, on the square lattice) corresponds to at most a sequence of BFACF moves on the original walk (in the cubic lattice). It is easy to prove the following proposition (since the BFACF algorithm is ergodic in the square lattice (Madras 1986)):

Proposition 2.2. Suppose that ω is a walk in the cubic lattice in W_{xy} with projection ν in the square lattice. Suppose that ν is a self-avoiding walk connecting the vertices u and v. Then there exist a sequence of BFACF moves on ω which will change the projection ν into any other self-avoiding walk connecting u and v.

3. Properties of the BFACF algorithm in three dimensions

In this section we consider the converse of proposition 2.1: if two polygons have the same knot type, then there exists a sequence of BFACF moves which connects them. That is, they are in the same ergodic class. An outline of the proof (of the converse of proposition 2.1) is as follows. We first show that if two polygons have identical lattice knot projections (as defined in section 2), then there exists a sequence of BFACF moves which will take one onto the other. If we know this, then all that is left to do is to prove that we can apply BFACF moves on the polygons so that they have identical lattice knot projection.

In order to accomplish this, we first show that we can apply BFACF moves to any given polygon to change its projection into a lattice knot projection. We can then restrict our attention to the set of polygons which have lattice knot projections. An ingredient in this proof is a construction which enables us to create as much space, empty of any images of edges or vertices, around a given point in the projection of the polygon, as we require. This construction is very useful in the execution of Reidemeister moves, since we can create as much 'working space' as needed to perform the required move at a given location in the projection. Therefore, we can take two polygons of the same knot type and with lattice knot projections, and change one of them by BFACF moves such that they have isotopic lattice knot projections. The last step in the proof is to show that we can make isotopic lattice knot projections identical.

In this section we shall need the following notation:

Definition 3.1. A *segment* $[\omega_i, \omega_j]$ of a polygon ω is the set of vertices ω_i, ω_{i+1}, ..., ω_j if $i \leqslant j$, or is the union of $[\omega_i, \omega_n]$ and $[\omega_0, \omega_j]$ if $j < i$.

Definition 3.2. A segment $[\omega_i, \omega_j]$ is a *side* if the union of all the edges associated with the vertices in the segment is a line segment. We denote a side connecting the vertices ω_i and ω_j by $[\omega_i, \omega_j]_s$.

A polygon in P_n has at least four, and and most n sides. (If the angles between a given edge and its nearest neighbour edges are 90°, then the segment itself is a side). We define a *side-operator*, which takes a side of a given polygon and shifts it in a desired direction. Let $[\omega_i, \omega_j]_s$ be any side in ω, and suppose that e_* is a unit vector perpendicular to the edges in $[\omega_i, \omega_j]_s$. The aim is to perform an operation on this side to transform it into the side $[\omega_i + e_*, \omega_j + e_*]_s$. This is easily done using BFACF operations: Consider first the edge $[\omega_i, \omega_{i+1}]$, and perform a BFACF move, whichever is necessary, to 'shift' this edge to $[\omega_i + e_*, \omega_{i+1} + e_*]$. Consider then, in succession, the edges $[\omega_{i+k}, \omega_{i+k+1}]$, until $i + k = j - 1$. We have then moved the side one unit distance in the e_* direction. Call this operation $\mathcal{M}_{ij}(e_*)$. Then we may write symbolically

$$\mathcal{M}_{ij}(e_*)[\omega_i, \omega_j]_s = [\omega_i + e_*, \omega_j + e_*]_s. \tag{3.1}$$

Note that we can use the operator $\mathcal{M}_{ij}(e_*)$ to split a side into sections, that is,

$$\begin{aligned}\mathcal{M}_{kl}(e_*)[\omega_i, \omega_j]_s = [\omega_i, \omega_k]_s \cup [\omega_k, \omega_k + e_*]_s \cup [\omega_k + e_*, \omega_l + e_*]_s \\ \cup [\omega_l + e_*, \omega_l]_s \cup [\omega_l, \omega_j]_s.\end{aligned} \tag{3.2}$$

where k and l are vertices on the side beween i and j. The inverse of the operator $\mathcal{M}_{ij}(e_*)$ is $\mathcal{M}_{ij}(-e_*)$.

We can now prove the following proposition on polygons with regular lattice knot projections (in the sense of section 2.2):

Proposition 3.3. Let ω and ν be two polygons in $\mathcal{Z}^3$ which have regular lattice knot projections. If ω and ν have identical lattice knot projections, then there exists a set of BFACF moves which connects ω and ν.

Proof. Since the projections are regular, they each consist of a set of line segments in the plane Q. Each of these line segments is the image of a number of sides in ω (or ν). Let $[\omega_i, \omega_j]$ be the segment of ω which has as image a particular line segment $\{i, j\}$ in the lattice knot projection of ω. If the line segment $\{i, j\}$ contains underpasses, then we argue as follows. Suppose that there is only one underpass at k. Separate $\{i, j\}$ into the pieces $\{i, k-1\} \cup \{k-1, k+1\} \cup \{k+1, j\}$. The segments $[\omega_i, \omega_{k-1}]$ and $[\omega_{k+1}, \omega_j]$ are then without underpasses, and we can apply the operator $\mathcal{M}_*(e_3)$ to the sides which are parallel to the plane Q. If there are more underpasses in the line segment, then we perform this operation at each of the underpasses. Since these segments have the same images in the lattice knot projections for both ω and ν, we can change the segments in ω into a standard position, and then change the segments in ν to be identical to these standard positions.

The polygons are now only different at the underpasses. By our construction, each of the underpasses is a segment which is a side of two edges. Apply the operator $\mathcal{M}_*(e_3)$ to each of these edges until it is unit distance from the vertex which overpasses it. The polygons will then also be identical at this underpass. Repeat this process at each underpass. Then ω and ν are identical. □

This proposition is important. It says that if we have two polygons, and if we can use BFACF moves to transform them into polygons which have identical regular lattice knot projections in the plane Q, then we can change them into identical polygons, using BFACF moves. We now show that we can transform any projection into a regular lattice knot projection and, if two regular lattice knot projections are of the same knot type, then we can make them identical.

Let ω be any polygon in $\mathcal{Z}^3$. Consider the projection of ω in the plane Q. The projection consists of a set of line segments joined at integer sites in Q. If this projection is not regular, then there is either at least one vertex between two line segments in the projection which is a multiple point in the projection, or there are two edges which are projected to the same edge in Q, or both. To proceed, we define the following:

Definition 3.4. Suppose that ω is a polygon. If every plane $\mathcal{T}$ which intersects the plane Q (the $(z = 0)$-plane) at 90° contains at most a finite number of disjoint points and at most *one segment* of ω, and if the projection of ω in Q contains no multiple edges, then we say that ω is in *standard form.*

A polygon in standard form has a regular lattice knot projection (by definition). Let $\mathcal{T}_i(a)$ be the plane perpendicular to the vector e_i containing the point a. We can now prove the following lemma:

Lemma 3.5. Let ω be a polygon. With $m \in \mathcal{Z}$, the plane $\mathcal{T}_i(m + \frac{1}{2})$ intersects ω in a finite number of points which are elements of a number of sides in ω. We can use BFACF moves to add an edge to each of the sides intersected by this plane.

Proof. Consider the plane through the maximum ith coordinate of the vertices in ω, say $\mathcal{T}_i(l)$. Label all the segments of ω in this plane by integers 1, 2, ..., and operate on each of them in turn by $\mathcal{M}_*(e_i)$. Once this is finished, consider the plane

$T_i(l-1)$, and so on, until we reach the plane $T_i(m+1)$. Since we have moved all the segments of ω out of this plane, the intersection of it with ω is at most a finite number of points which are elements in the sides of ω which penetrate the plane $T_i(m+\frac{1}{2})$. Moreover, these sides terminate in segments which have been moved by the operator $\mathcal{M}_*(e_i)$, so they are each unit distance longer. □

We can now consider the effect of lemma 3.5 on the projection of a polygon. The image of the plane $T_i(m+\frac{1}{2})$ is a line in Q. Using the construction, we can 'move' every edge in the polygon which projects to one side of the line by one lattice distance. The effect is that we 'open' up space in the projection; the line $T_i(m+1)$ will contain no edges of the projected polygon. By repeating this process we can create as much space as we want. We can now consider the application of lemma 3.5 to projections. In the next proposition we show that we can use BFACF moves to change any polygon into a polygon with a standard projection in the plane Q.

Proposition 3.6. Let ω be any polygon. Then we can apply BFACF moves to ω to transform it into a polygon in standard form.

Proof. Any plane which intersects the plane Q at $90°$ and which is not perpendicular to either the unit vectors e_1 or e_2 will intersect ω in finitely many points. Therefore, we have only to consider the intersections between ω and planes $T_i(a)$ perpendicular to either e_1 $(i=1)$ or e_2 $(i=2)$, and therefore perpendicular to the plane Q. Without loss of generality, let us first probe ω with $T_1(a)$. Let $X(\omega_i)$, $Y(\omega_i)$ and $Z(\omega_i)$ be the components of the vertex ω_i in the e_1, e_2 and e_3 directions respectively.

Probe ω with $T_1(\max_i X(\omega_i))$, which is the plane through the maximum value of the e_1 component of vertices in ω. The intersection between this plane and ω is a set of segments of ω. We can choose any of these segments and operate with $\mathcal{M}_*(e_1)$ on the sides which make up this segment. This will move the segment one unit distance in the e_1 direction, and reduce the number of segments in the intersection by one. Label the shifted segment by the integer 1. If there is still more than one segment left in the intersection, then we apply $\mathcal{M}_*(e_1)$ first to the segment labelled by a 1, and then to any of the other segments, which we label by a 2. At the kth step, we first apply $\mathcal{M}_*(e_1)$ to all the segments labelled by $1, 2, \ldots, k-1$, and then to the kth segment. Eventually, there will be only one segment in the intersection between ω and $T_1(\max_i X(\omega_i))$, and each of the planes parallel to this plane through the points $\max_i X(\omega_i)+l$, l an integer, contains at most one segment. In this case we label the last segment by the next integer, say m, and consider the plane $T_1(\max_i X(\omega_i)-1)$, and continue the process, while we label the new segments by $m+1$, $m+2$, and so on.

Since the polygon is finite this process must end. There will then be at most one segment in the intersection of $T_1(\max_i X(\omega_i))$ with ω. So far, we have not dealt with multiple edges in the projection of ω. Let $[\omega_i, \omega_j]$ be a segment in a plane $T_1(a)$ such that its projection *contains multiple edges*. The segment contains several sides, and some the edges in these sides are projected to the same edges in Q. Choose k such that $[\omega_i, \omega_k]$ is the biggest subsegment which has no multiple edges in its projection. Apply the construction in lemma 3.5 to remove all segments of the polygon from the plane $T_1(a+1)$. Then we can apply the operator $\mathcal{M}_{ik}(e_1)$ to $[\omega_i, \omega_k]$. The newly created segment in plane $T_1(a+1)$ contains no double edges in its projection, and we have reduced the number of edges in the segment in plane $T_1(a)$ by at least one. We repeat this process now for the segment $[\omega_k, \omega_j]$. At every stage we move a segment with no double edges in its projection to the plane $\mathcal{T}_1(a+1)$. Since we started with a finite number of edges in $[\omega_i, \omega_j]$, this process must stop after a finite number of steps. The newly created segments have no double edges in their projections, and we have removed the multiple edges from the projection of the segment.

At this stage we have 'stretched' ω in the e_1 direction such that no edge in the e_2 direction is mapped into a multiple edge, and each plane $\mathcal{T}_1(a)$ contains at most one segment of ω. The polygon is not yet in standard form. The edges in the e_1 direction form segments which may be more than one per plane, or may map to multiple edges in Q. We now rotate our axes, and probe ω with the planes $\mathcal{T}_2(a)$ in the same manner as before, with one important difference: we note that every application of $\mathcal{M}_*(e_2)$ in this part of the construction introduces or deletes edges in ω in the e_2 direction, and we may create a second segment in one of the planes $\mathcal{T}_1(a)$. If this happens, then we first apply lemma 3.5 to create an empty plane for the new edges.

As before, this process must stop, since ω is finite. By construction, ω is now in standard form. □

By definition 3.4 and proposition 3.6 we have:

Corollary 3.7. Let ω be any polygon. Then, by applying BFACF moves, we can transform ω into a polygon which has a regular lattice knot projection.

We now show that we can transform two polygons of the same knot type, each with a regular lattice knot projection, into polygons with the same lattice knot projection.

Let ω and ν be two lattice polygons of the same knot type, and suppose that both have regular lattice knot projections on the plane Q. If we forget about the lattice for the moment, then there is a sequence of Reidemeister moves which connects a plane isotopy of the projection of ω to a plane isotopy of ν. The next step is to consider the execution of Reidemeister moves on a regular lattice knot projection. In these constructions we pass one segment of the polygon over another (in the projection). Observe that we can always do this; if we perform BFACF moves in the e_3 direction, we can increase the e_3 component of the vertices of the overpassing strand at will, without changing the projection. Consider the moves one by one:

Reidemeister I. Consider this situation in figures 2 and 3. This move operates on a segment of ω which starts in an overpass (or underpass) and ends in an underpass (or overpass) at the same location in the lattice Q, without containing any crossings between its beginning and its end. Let this segment be $[\omega_i, \omega_j]$. Arrange matters such that $i < j$, without loss of generality. This segment consists of a number of sides. Operate with $\mathcal{M}_*(\pm e_3)$ on all the sides parallel to the plane Q so that the third components of all the sides are now the same as that of ω_i, except for the side $[\omega_{j-1}, \omega_j]_s$. The segment $[\omega_i, \omega_{j-2}]$ is then planar, and since the BFACF algorithm is ergodic in two dimensions, we can operate on this segment until it consists only of the edge $[\omega_i, \omega_{j-2}]$. Alternatively, we can simply apply proposition 2.2 and the operator $\mathcal{M}_*(\pm e_3)$ to the segment $[\omega_{i+1}, \omega_{j-1}]$. One more application of a BFACF move to reduce the length by 2 removes the extra pair of edges, completing the construction. The sequence of events is illustrated in figure 3. In the opposite direction we must include a segment into ω which will change its projection by Reidemeister I. In this case we reverse the last two steps of the previous construction. To do this, we must create an area in Q, free of any projected edges, using the constructions in lemma

3.5. Note that the projection of ω is still a regular lattice knot projection; we have to apply lemma 3.6 enough times to prevent the formation of double edges, or of more than one segment in each plane $\mathcal{T}_i(z)$, $i = 1, 2$.

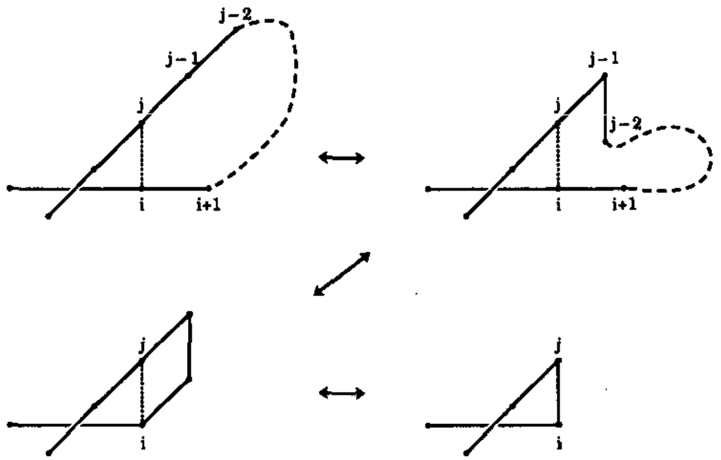

Figure 3. A Reidemeister move of type I performed on a polygon.

Reidemeister II. Consider this situation in figures 2 and 4. This move operates on two segments of ω which start at a crossing, and end in a second crossing, without any other crossings in the two segments. Let these segments be $[\omega_i, \omega_j]$ and $[\omega_k, \omega_l]$. To perform the move, we create an open area in Q, free of projected edges. The area P in figure 4 can be made arbitrarily large in any direction, using the construction in lemma 3.5. We can then operate on the segments involved in the move. We first make them planar by applying $\mathcal{M}_*(\pm e_3)$, and then perform the move by applying BFACF moves, or alternatively, apply proposition 2.2. The resulting projection is still a regular lattice knot projection, since we can create enough space in Q to prevent the appearances of double edges, and keep the number of segments in any plane $\mathcal{T}_i(z)$ at most one. The opposite of this move is now obvious.

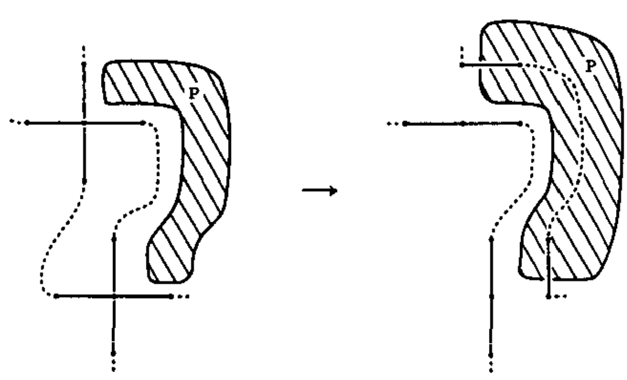

Figure 4. Reidemeister II performed on a projection of a polygon.

Reidemeister III. The last Reidemeister move is also straightforward. We illustrate it in figures 2 and 5. Once again, we apply lemma 3.5 to make the areas P_1, P_2 and P_3 arbitrarily large. The move can then be performed by making a segment of the polygon planar and applying BFACF moves (or use proposition 2.2), as for Reidemeister I and II. The opposite of this move is now obvious.

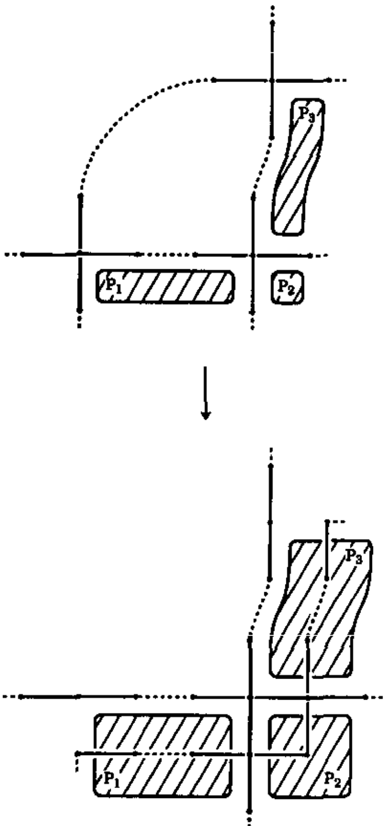

Figure 5. Reidemeister III performed on a projection of a polygon.

We therefore have the following proposition:

Proposition 3.8. Let ω and ν be polygons of the same knot types, each with a regular lattice knot projection, W and V. Then, by applying BFACF moves, we can transform ω into ω', such that there exists a plane isotopy $\mathcal{I}$ with $\mathcal{I} : (W', Q) \rightarrow (V, Q)$, where W' is the regular lattice knot projection of ω'.

Proof. If we consider W and V as two knot projections in Q, then there exists a sequence of Reidemeister moves connecting W to V. We can now perform these Reidemeister moves one by one, by performing BFACF moves on ω, applying lemma 3.7 and the constructions for the moves as set out above. Finally, (W', Q) will be isotopic to (V, Q). □

There is now just one last step to check. We must be able to make identical two isotopic projections through BFACF moves. This involve two phases: the first is a comparison of the two projections, edge by edge, and the second is a possible 90° rotation or a translation of the projection.

Proposition 3.9. Let ω and ν be two polygons with regular lattice knot projections W and V such that there exists an isotopy of the plane $\mathcal{I} : (W, Q) \to (V, Q)$. Then we can perform BFACF moves such that $\mathcal{I}$ is the identity or a translation of the lattice $\mathcal{Z}^2$.

Proof. Since the projection W is a regular projection, it does not contain any double edges in Q, and the only double points are at over- and underpasses, where sides make 90° angles with each other. By the Jordan curve theorem, W separates Q into at least two components, at most one of which is infinite. Let $d(x, y) = |y_1 - x_1| + |y_2 - x_2|$ be the distance between the vertices $x = (x_1, x_2)$ and $y = (y_1, y_2)$ in Q. Let $\{w_i\}$ and $\{v_i\}$ be the crossings in the projections W and V respectively. Apply lemma 3.5 to ω until every area in V can be covered by its corresponding area in W, and until the minimum distance between a pair of crossings in W is greater than the maximum distance between pairs of *vertices* in V.

If two areas in a regular lattice knot projection share a projected segment of the polygon (in Q), then we say that these areas are *nearest neighbours*. Label any of the areas in W by the integer 1. By virtue of the isotopy $\mathcal{I}$, there is an area in V which corresponds to this area. Call these areas W_1 and V_1. Superimpose W_1 and V_1 such that V_1 is covered by W_1. By our construction, this is always possible. W_1 is bounded by a number of projected segments of ω. Fix the endpoints of these projected segments (the vertices t in figure 6), and apply proposition 2.2 to 'close' W_1 in on V_1. (Note that we can perform BFACF moves on the projection to take the area $W_1 - V_1$ to zero; by proposition 2.2 we can always do this). Then every projected segment bounding V_1 is in the image of the corresponding segment (under the isotopy) in W. W_1 is then identical to V_1. We illustrate this in figure 6. In this construction we may introduce double edges in the projection W, as we see in figure 6. We leave this unchanged for the moment. Note that the vertices in W which are incident on W_1 have been moved to their 'correct position' by this operation.

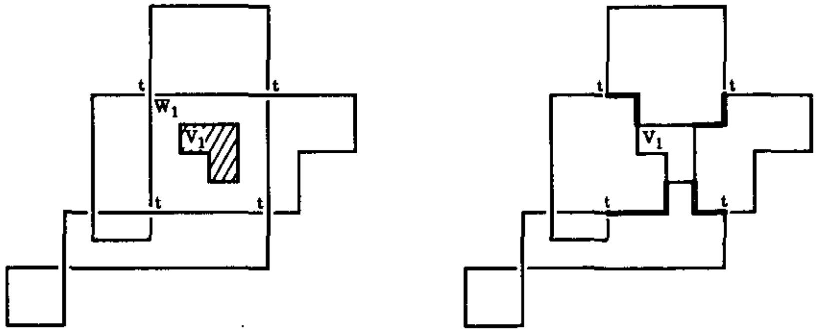

Figure 6. To make W_1 identical to V_1 we fix the segments bounding W_1 at the vertices t and apply proposition 2.2.

Consider all the nearest neighbours of W_1, and label them, in any order, 2, 3, ... We repeat the process with W_2, starting with the segments which share a vertex with a labelled area. We have to consider two possibilities. The first case is the same as the above; V_2 is covered by W_2. In this case we perform the same construction; we 'close' the relevent segments in on V_2 to make the area $W_2 - V_2$ zero. By proposition 2.2 we can always do this. The second possibility is that there may be some segments in W which pass through the image of V_2. In this case we must sweep V_2 clean before we can apply the construction. But we can always do that by proposition 2.2. Note that the segments of W that we move cannot interact with any crossing in W; if they do, then there is a segment in V which can connect two crossings in W. This is a contradiction. Once we have swept all the segments out of V_2, then we can apply the above construction to make W_2 and V_2 identical. W_2 shares a segment with W_1, so we do not change this segment since it is already in the correct conformation. Note that we never perform a Reidemeister move in these operations; the worst we do is to introduce multiple edges in strings at the crossings in W. Once we have performed the construction at W_2, we consider W_3, until all the nearest neighbours of W_1 have been changed to be identical to the corresponding areas in V. W_1 and its neighbours form a subset of W which is identical to the corresponding areas in V. Then consider nearest neighbours of W_2, W_3, and so on, until no more neighbours can be found. If all the segments which bound an unlabelled area are also parts of the boundaries of labelled areas then the unlabelled area will be identical to the corresponding area in V. We label these areas with an A, since we do not need to label them with an integer. One such event occurs in figure 6. We call this set of areas a *cluster* around the area W_1. Any unlabelled area in W cannot be a neighbour of a labelled area, and is connected to the cluster around W_1 either by a crossing, or by a string of double edges which was created in the construction.

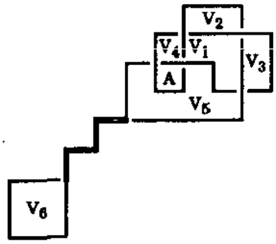

Figure 7. The areas labelled V_i, $i = 1$ to 5, and A form a cluster. V_6 forms a second cluster.

Choose any unlabelled area in W which is connected to a labelled cluster by either a crossing, or by a string of double edges. Consider the vertex t (figure 7), incident on the labelled cluster, and common to the chosen area, or to the string of double edges connected to the labelled area. The two segments bounding the unlabelled area and incident on t in the projection contain the string of double edges, if such a string exists. If we now perform BFACF moves to identify these segment to their counterparts in V, then the double edges disappear, since they do not exist in V. We can now operate on the other segments to make the area identical to its image in V. Label this area by k. Consider now the nearest neighbours of W_k until a second cluster has been identified. Look then for a third cluster, and so on, until all the areas are labelled.

Once every area in W is labelled, then W is identical to V up to an accidental translation of the lattice $Z^2 \subset Q$ (which was introduced by the initial position of the area W_1). $\square$

Proposition 3.10. Let ω be any polygon with projection W in the plane Q. Then we can apply BFACF moves to translate W any distance in a lattice direction in Q.

Proof. Without loss of generality, suppose we want to translate W a distance d along the e_1 direction. Let us probe the polygon by $\mathcal{T}_1(z)$, where we choose $z = \max_i(X(\omega_i))$. $\mathcal{T}_1(z)$ contains a number of sides of ω. Pick any of these sides, say $[\omega_i, \omega_j]_s$ and apply the operation $\mathcal{M}_{ij}(e_1)$ d times in succession. This will translate $[\omega_i, \omega_j]_s$ a distance d in the e_1 direction. Repeat this process with all the sides in $\mathcal{T}_1(z)$. Reduce z then by 1 and repeat the process. Keep on reducing z until we cannot find an intersection between ω and $\mathcal{T}_1(z)$. Then we have translated W a distance d in the e_1 direction. □

We have now completed the proof of the following theorem.

Theorem. 3.11 The ergodicity classes of the BFACF algorithm, when applied to unrooted polygons, are the knot types of the polygons.

The proof of theorem 3.11 is direct from proposition 3.3, corollary 3.7, and propositions 3.8, 3.9 and 3.10.

4. Discussion

The application of the BFACF algorithm to self-avoiding polygons on the simple cubic lattice involves sampling along a realization of a Markov chain defined on the set of polygons. We have shown that the ergodic classes of this Markov chain are the knot classes of the polygons. This means that the set of polygons sampled in any realization is the set which has the same knot type as the initial polygon. This makes the BFACF method a convenient one for studying the properties of polygons with given knot type.

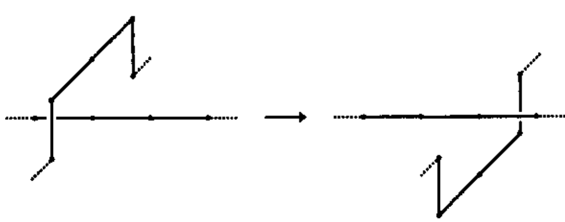

Figure 8. A crankshaft which 'passes' one segment of a polygon through another.

In two dimensions the corresponding algorithm is known to be ergodic (Madras 1986). In four and higher dimensions the question is still open though we think it unlikely that topological obstructions will occur.

A question which arises naturally in three dimensions is what extra moves are needed to make the algorithm ergodic? Brower (1991) has suggested that the addition of a four-bond crankshaft move to the BFACF algorithm will make the algorithm ergodic for all polygons. To see this, suppose we have a knotted polygon ω. By applying BFACF moves we can put ω into any convenient conformation. In particular, we can transform ω to a polygon whose vertices have third-component values only 0 or 1. In addition to this, we can arrange matters such that the only vertices in ω with third components not zero are those involved in an overpass. (Think of a knot projection, where the vertices are all in the same plane, except at the crossings, where the overpasses are like the conformations in figure 8). Since any knot universe has an associated set of under- and overcrossings which corresponds to the unknot (e.g. the knot with ascending overcrossings) we can untie any knot by performing the crankshaft move, as we illustrate in figure 8, in turn at every crossing which has the undesired orientation.

Acknowledgments

The authors acknowledge helpful discussions with H Lim, N Madras, A D Sokal and D W Sumners. This research was supported by a grant from NSERC of Canada. This work was also partially supported by the US Department of Energy under Contract no DE-FC05-85ER250000.

References

Aragao de Carvalho C and Caracciolo S 1983 *J. Physique* **44** 323
Aragao de Carvalho C, Caracciolo S and Frölich J 1983 *Nucl. Phys.* B **251** 209
Berg B and Foester D 1981 *Phys. Lett.* **106B** 323
Brower R 1991 Unpublished
Bluestone S and Vold M J 1965 *J. Chem. Phys.* **42** 4175
Burde G and Zieschang H 1985 *Knots* (Berlin: de Gruyter)
Caracciolo S, Pelissetto A and Sokal A D 1990a *J. Stat. Phys.* **60** 1
Caracciolo S, Pelissetto A and Sokal A D 1990b *J. Phys. A: Math. Gen.* **23** 4509
Caracciolo S, Pelissetto A and Sokal A D 1991 *J. Stat. Phys.* **63** in press
Caracciolo S and Sokal A D 1986 *J. Phys. A: Math. Gen.* **19** L797
Hammersley J M and Handscomb D C 1964 *Monte Carlo Methods* (London: Methuen)
Janse van Rensburg E J and Whittington S G 1991 *J. Phys. A: Math. Gen.* **24** 3935
Kemeny J G and Snell J L 1976 *Finite Markov Chains* (Berlin: Springer)
Kron A K 1965 *Vysokomol. Soedin.* **7** 1228
Lal M 1969 *Mol. Phys.* **17** 57
Madras N and Slade G 1992 *The Self-avoiding Walk* (Boston: Birkhauser) in preparation
Madras N and Sokal A D 1987 *J. Stat. Phys.* **47** 573
Metropolis M, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 *J. Chem. Phys.* **21** 1087
Monnerie L and Geny F 1969 *J. Chim. Phys.* **66** 1691
Reidemeister K 1932 *Knotentheorie* (Berlin: Springer)
Sokal A D and Thomas L E 1988 *J. Stat. Phys.* **50** 109
Verdier P H and Stockmayer W H 1962 *J. Chem. Phys.* **36** 227