# Irreducibility of the Tutte polynomial of a connected matroid

C. Merino [1] A. de Mier [2] M. Noy [2]

**Abstract**

We solve in the affirmative a conjecture of Brylawski, namely that the Tutte polynomial of a connected matroid is irreducible over the integers.

If $M$ is a matroid over a set $E$, then its Tutte polynomial is defined as

$$T(M;x,y) = \sum_{A \subseteq E} (x-1)^{r(E)-r(A)}(y-1)^{|A|-r(A)},$$

where $r(A)$ is the rank of $A$ in $M$. This polynomial is an important invariant as it contains much information on the matroid; see [2, 3] for useful surveys.

One of the basic properties of $T(M;x,y)$ is that, if $M$ is the direct sum of two matroids $M_1$ and $M_2$, then

$$T(M;x,y) = T(M_1;x,y)\,T(M_2;x,y).$$

In particular, this implies that $T(M;x,y)$ has a non-trivial factor in $\mathbb{Z}[x,y]$ if $M$ is disconnected. Brylawski [1] conjectured that the converse also holds; this paper is devoted to a proof of this conjecture.

**Theorem 1** *If $M$ is a connected matroid, then $T(M;x,y)$ is irreducible in $\mathbb{Z}[x,y]$.*

Actually, an analysis of our proof shows that $T(M;x,y)$ is irreducible even in $\mathbb{C}[x,y]$.

**Corollary 1** *If a matroid $M$ has $c$ connected components $M_1,\ldots,M_c$, then the factorization of $T(M;x,y)$ in $\mathbb{Z}[x,y]$ is exactly*

$$T(M;x,y) = T(M_1;x,y)\,\cdots\,T(M_c;x,y).$$

The main tool in proving our result is the following set of linear equations $(B_k)$ that are satisfied by the coefficients of the Tutte polynomial of any matroid, and that were proved in [1].

**Lemma 1** *Let $T(M;x,y) = \sum b_{ij}x^iy^j$ be the Tutte polynomial of a matroid $M$ and let $m$ be the number of elements in $M$. Then*

$$(B_k) \qquad \sum_{s=0}^{k}\sum_{t=0}^{k-s}(-1)^t\binom{k-s}{t}b_{st} = 0,$$

*for $k = 0,1,\ldots,m-1$.*

[1] Instituto de Matemáticas, U.N.A.M., Mexico. This author was supported by a grant from D.G.A.P.A.

[2] Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya, Barcelona, Spain. Partially supported by projects SEUI-PB98-0933 and by CUR Gen. Cat. 1999SGR00356.

We also need the following basic properties [2]:

1. $b_{00} = 0$ if $|E| \geq 1$;
2. $b_{10} \neq 0$ if and only if $M$ is connected;
3. $x^k \,|\, T(M;x,y)$ if and only if $M$ has at least $k$ coloops;
4. $y^k \,|\, T(M;x,y)$ if and only if $M$ has at least $k$ loops;
5. if $i \geq r(M)$ or $j \geq n(M)$, then $b_{ij} = 0$, except if $i = r(M)$ and $j = 0$, or if $i = 0$ and $j = n(M)$, where $r(M)$ is the rank of $M$ and $n(M) = m - r(M)$ is the nullity of $M$.

Suppose now $M$ is a connected matroid on a set of $m$ elements, and that there is a non-trivial factorization

$$T(M;x,y) = \sum b_{ij}x^iy^j = A(x,y)\,C(x,y), \tag{1}$$

where $A(x,y) = \sum a_{ij}x^iy^j$ and $C(x,y) = \sum c_{ij}x^iy^j$.

Since $b_{00} = 0$, either $a_{00}$ or $c_{00}$ is zero; we may assume $a_{00} = 0$. Since

$$0 \neq b_{10} = a_{00}c_{10} + a_{10}c_{00},$$

th! assumption implies that $c_{00} \neq 0$. We will prove that $c_{00} = 0$, thus obtaining a contradiction.

Since $M$ is connected, by properties 3 and 4 above, neither $x$ nor $y$ are factors of $A(x,y)$ or $C(x,y)$. Define for a polynomial $P(x,y) = \sum p_{ij}x^iy^j$,

$$r_P(x) = \max\{i \,:\, p_{i0} \neq 0\}, \qquad r_P(y) = \max\{j \,:\, p_{0j} \neq 0\}, \tag{2}$$

and let

$$m(P) = r_P(x) + r_P(y).$$

Clearly, from (1), $m = m(T) = m(A) + m(C)$. As we are supposing a non-trivial factorization, it follows that $r_A(x), r_A(y) \leq m(A) < m(T)$. Also useful is the following property.

**Lemma 2** *Let $M$ be a connected matroid and $T(M;x,y)$ be its Tutte polynomial with a factorization as in* (1). *Then the polynomial $A(x,y)$ also satisfies property 5, that is, if $r_A(x) \leq i$ or $r_A(y) \leq j$ then $a_{ij} = 0$, except if $i = r_A(x)$ and $j = 0$, or if $i = 0$ and $j = r_A(y)$.*

*Proof.* Let $\alpha = \max\{i : a_{ij} \neq 0 \text{ for some } j\}$ and $\beta = \max\{j : a_{\alpha j} \neq 0\}$; define analogously $\alpha'$ and $\beta'$ for the polynomial $C(x,y)$. The monomial $a_{\alpha\beta}c_{\alpha'\beta'}x^{\alpha+\alpha'}y^{\beta+\beta'}$ appears in $T(M;x,y)$, as it cannot be cancelled, and it is the term with maximum degree of $x$ in $T(M;x,y)$. Using property 5 we see that $\alpha + \alpha' = r(M)$ and $\beta + \beta' = 0$, so $\beta = 0$ and $\alpha = r_A(x)$. Thus, the maximum degree of $x$ in $A(x,y)$ has coefficient $a_{r_A(x),0}$. A similar argument shows that the maximum degree of $y$ in $A(x,y)$ has coefficient $a_{0,r_A(y)}$. □

We next prove two lemmas that together imply Theorem 1.

Let $(A_k)$, $k = 0, 1, \ldots, m(A)$ be the same set of equations as the $(B_k)$, but with the $a_{st}$ replacing the $b_{st}$, that is,

$$\sum_{s=0}^{k}\sum_{t=0}^{k-s}(-1)^t\binom{k-s}{t}a_{st} = 0. \tag{$A_k$}$$

Note that we do *not* assume $A(x,y)$ to be the Tutte polynomial of a matroid, hence we do not know whether equations $(A_k)$ hold or not. In fact, we have the following result.

**Lemma 3** *With hypothesis as in Lemma 2, there is at least one equation* $(A_l)$ *with* $r_A(x) \leq l \leq m(A)$ *that does not hold.*

*Proof.* First, for $r_A(x) \leq k \leq m(A)$ and $i \geq 0$ we define the equation $(A_{ki})$ as follows:

$$\sum_{s=0}^{k} \sum_{t=0}^{k-s} (-1)^{t+i} \binom{k-s}{t} a_{s,t+i} = 0. \tag{$A_{ki}$}$$

Note that $(A_{k0})$ is the same equation as $(A_k)$. Now we prove a recurrence relation involving these equations.

Observe that for $i > 0$ and $k > r_A(x)$ the left-hand side of equation $(A_{k,i-1})$ is:

$$\sum_{s=0}^{k} \sum_{t=0}^{k-s} (-1)^{t+i-1} \binom{k-s}{t} a_{s,t+i-1}. \tag{3}$$

Using the fact that $\binom{k-s}{t} = \binom{k-s-1}{t} + \binom{k-s-1}{t-1}$, and assuming $\binom{a}{-b} = 0$ for $a \geq 0$, $b > 0$, and also $\binom{a}{b} = 0$ if $a < b$, we can rewrite (3) in the following way:

$$\begin{aligned}
&\sum_{s=0}^{k} \sum_{t=0}^{k-s} (-1)^{t+i-1} \left[ \binom{k-1-s}{t} + \binom{k-1-s}{t-1} \right] a_{s,t+i-1} + (-1)^{i-1} a_{k,i-1} = \\
&\sum_{s=0}^{k-1} \sum_{t=0}^{k-1-s} (-1)^{t+i-1} \binom{k-1-s}{t} a_{s,t+i-1} + \sum_{s=0}^{k-1} \sum_{t=1}^{k-s} (-1)^{t+i-1} \binom{k-1-s}{t-1} a_{s,t+i-1} \\
&+ (-1)^{i-1} a_{k,i-1}.
\end{aligned}$$

The last term appears because $\binom{0}{0}$ cannot be decomposed into two binomial coefficients. But by Lemma 2 for this last term we have $a_{k,i-1} = 0$, as $r_A(x) < k$ . Also, the first and second terms in the second row +f the last expression (after a change of variables) are, respectively, the left-hand side of equations $(A_{k-1,i-1})$ and $(A_{k-1,i})$. So we can write symbolically:

$$(A_{k,i-1}) = (A_{k-1,i}) + (A_{k-1,i-1})$$

or

$$(A_{k-1,i}) = (A_{k,i-1}) - (A_{k-1,i-1}) \tag{4}$$

for $r_A(x) < k \leq m(A)$ and $i > 0$.

Let us suppose now that all equations $(A_k)$ hold for $r_A(x) \leq k \leq m(A)$ and we will find a contradiction. Consider equation $(A_{r_A(x),r_A(y)})$. By Lemma 2, the only term $a_{ij}$ involved in this equation that is not zero is $a_{0,r_A(y)}$. Then the left-hand side of $(A_{r_A(x),r_A(y)})$ reduces to $\binom{r_A(x)}{0} a_{0,r_A(y)} = a_{0,r_A(y)}$ that is different from zero. On the other hand, using Equation (4) repeatedly $r_A(y)$ times, we can express this nonzero term as a sum of the left-hand sides of equations $(A_{k0})$ for $r_A(x) \leq k \leq m(A) = r_A(x) + r_A(y)$, that we are assuming to be all equal to zero. Therefore we obtain a contradiction and we conclude that not all of the $(A_{k0}) = (A_k)$ hold for $r_A(x) \leq k \leq m(A)$. □

**Lemma 4** *If the coefficients* $a_{ij}$ *do not satisfy equation* $(A_k)$ *for some* $k \leq m(A)$*, then* $c_{00} = 0$.

*Proof.* Let $(A_k)$ be the first equation that does not hold. Equation $(B_k)$ holds because $k \leq m(A) < m$. First, we rewrite this equation taking into account that

$$b_{st} = \sum_{h \leq s} \sum_{l \leq t} c_{hl} a_{s-h,t-l}.$$

Then we have the following equalities for the left-hand side of $(B_k)$.

$$\sum_{s=0}^{k}\sum_{t=0}^{k-s}(-1)^t\binom{k-s}{t}b_{st} = \sum_{s=0}^{k}\sum_{t=0}^{k-s}(-1)^t\binom{k-s}{t}\sum_{h\le s}\sum_{l\le t} c_{hl}a_{s-h,t-l} =$$

$$c_{00}\sum_{s=0}^{k}\sum_{t=0}^{k-s}(-1)^t\binom{k-s}{t}a_{st} + \sum_{0<h+l\le k} c_{hl}\left[\sum_{s=h}^{k}\sum_{t=l}^{k-s}(-1)^t\binom{k-s}{t}a_{s-h,t-l}\right]. \tag{5}$$

Note that each $c_{hl}$ has as coefficient an expression similar to the left hand side $(A_k)$; in particular, for $c_{00}$ this coefficient is exactly the left-hand side of equation $(A_k)$. More precisely, we introduce the equation $(A'_{ni})$.

$$\sum_{s=0}^{n-i}\sum_{t=0}^{n-i-s}(-1)^{t+i}\binom{n-s}{t+i}a_{st} = 0. \tag{$A'_{ni}$}$$

Observe that the left-hand side of equation $(A'_{k-h,l})$ is the coefficient of $c_{hl}$ in (5) above: change indices $s \leftarrow s+h, t \leftarrow t+l$, and remark that for $s > k-l$ and $t \geq l$ the binomial $\binom{k-s}{t}$ vanishes. Also note that $(A'_{n0})$ is precisely equation $(A_n)$, which we are assuming holds for $0 \leq n < k$. Now, we prove that $(A'_{ni})$ holds for $1 \leq n \leq k$ and $1 \leq i \leq n$ using induction on $n$.

If $n = 1$, the only possible value for $i$ is 1 and $(A'_{11})$ reduces to $a_{00} = 0$, which was supposed from the beginning. Assuming the result for all values less than $n$, we use again a formula for the binomial coefficients to decompose the left-hand side of equation $(A'_{ni})$ into a sum of previous equations:

$$\sum_{s=0}^{n-i}\sum_{t=0}^{n-i-s}(-1)^{t+i}\binom{n-s}{t+i}a_{st} =$$
$$\sum_{s=0}^{n-i}\sum_{t=0}^{n-i-s}(-1)^{t+i}\left[\binom{n-s-1}{t+i-1}+\binom{n-s-2}{t+i-1}+\cdots+\binom{t+i-1}{t+i-1}\right]a_{st}.$$

Each binomial coefficient $\binom{n-s}{t+i}$ is partitioned into exactly $n-s-t-i+1$ terms, so that the last expression equals

$$\sum_{s=0}^{n-1-(i-1)}\sum_{t=0}^{n-1-(i-1)-s}(-1)^{t+i}\binom{n-1-s}{t+i-1}a_{st} +$$
$$+\sum_{s=0}^{n-2-(i-1)}\sum_{t=0}^{n-2-(i-1)-s}(-1)^{t+i}\binom{n-2-s}{t+i-1}a_{st} + \cdots + \sum_{s=0}^{0}\sum_{t=0}^{0}(-1)^{t+i}\binom{t+i-1}{t+i-1}a_{st}.$$

Now it is easy to check that the $p$-th term in the last sum is equal to the left hand side of equation $(A'_{n-p,i-1})$, for $1 \leq p \leq n-i+1$. Thus we obtain the following relation:

$$(A'_{ni}) = -\left((A'_{n-1,i-1}) + (A'_{n-2,i-1}) + \cdots + (A'_{i-1,i-1})\right).$$

If $i = 1$, the equations on the right are $(A_{n-1}),\ldots, (A_0)$, all of which hold because $n-1 < k$. If $i > 1$, equations $(A'_{n-1,i-1}),\ldots, (A'_{i-1,i-1})$ hold by inductive hypothesis. In both cases $(A'_{ni})$ holds, and this concludes the induction.

Using this result we see from (5) that equation $(B_k)$ reduces to $c_{00}(A_k) = 0$. As $(A_k)$ does not hold, $c_{00}$ must be zero and the lemma is proved. □

The above two lemmas show that $c_{00} = 0$ and this establishes the theorem.

**Remark.** The assumption of characteristic zero is necessary, since otherwise property 2 after Lemma 1 does not hold, that is, $b_{10}$ can be zero because of the characteristic. For example,

$$\begin{aligned} T(M(K_4); x, y) &= 2x + 2y + 3x^2 + 4xy + 3y^2 + x^3 + y^3 \\ &= (x + y)(x + y + x^2 + xy + y^2) \mod 2, \end{aligned}$$

whereas $M(K_4)$ is a connected matroid.

## References


[1] T. Brylawski, A decomposition for combinatorial geometries, *Trans. Amer. Math. Soc.* 171 (1972), 235–282.

[2] T. Brylawski and J. Oxley, The Tutte Polynomial and Its Applications, in *Matroid Applications* (ed. N. White), Cambridge Univ. Press (1993).

[3] D. Welsh, *Complexity: knots, colourings and counting*, Cambridge Univ. Press (1993).