Kastelyn's Formula for Perfect Matchings

Rich Schwartz

April 8, 2013

1 Pfaffian Orientations on Planar Graphs

The purpose of these notes is to prove Kastelyn's formula for the number of perfect matchings in a planar graph. For ease of exposition, I'll prove the formula in the case when G has the following description: G can be drawn inside a polygon P such that $P - G$ is a finite union of smaller polygons. I'll call such a graph *nice*. See Figure 1.

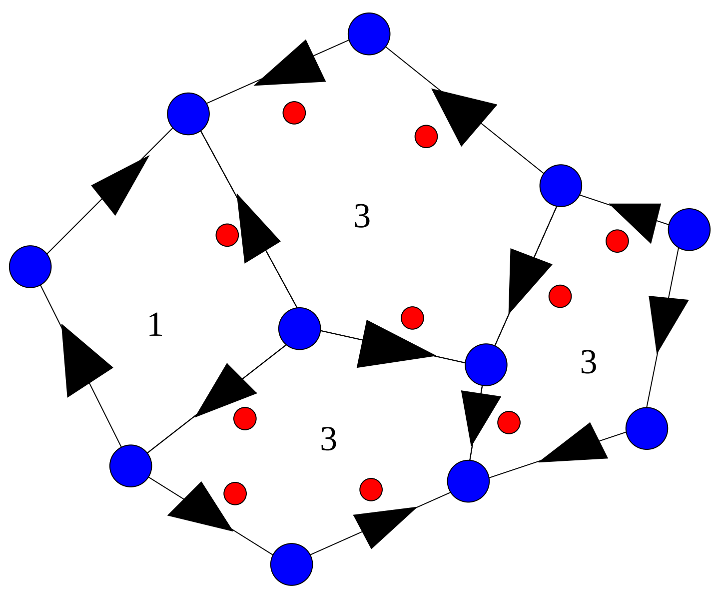

Figure 1: a nice planar graph

An *elementary cycle* of G is a cycle that does not surround any other vertices. In Figure 1, there are 4 elementary cycles. A *Pfaffian orientation* on G is a choice of orientations for each edge such that there are an odd number of clockwise edges on each elementary cycle. Figure 1 shows a Pfaffian orientation. The red dots indicate the clockwise edges, with respect to each elementary cycle.

Lemma 1.1 *Any nice graph has a Pfaffian orientation.*

Proof: Choose a spanning tree T for G and orient the edges of T arbitrarily. Define a *dual graph* as follows. Place one vertex in the interior of each elementary cycle of G, and a root vertex for the outside of the big polygon, and then connect the vertices by an edge if they cross an edge not in T. Call this dual graph T^*. If T^* is not connected, then T contains a cycle. Since T contains a cycle, T^* is connected. If T^* contains a cycle, then T is not connected. Since T is connected, T^* contains no cycle. In short T^* is a tree.

Let v be any node of T^*. The vertex v joins to an edge w of T^* across an edge e^*. The edge e^* crosses one edge e of the corresponding elementary cycle of G. Orient e so that this cycle has an odd number of clockwise oriented edges, then cross off v and e^*, leaving a smaller tree. Now repeat, using a node of the smaller tree. And so on. At each step, there is one way to orient the edge to make things work for the current elementary cycle. ♠

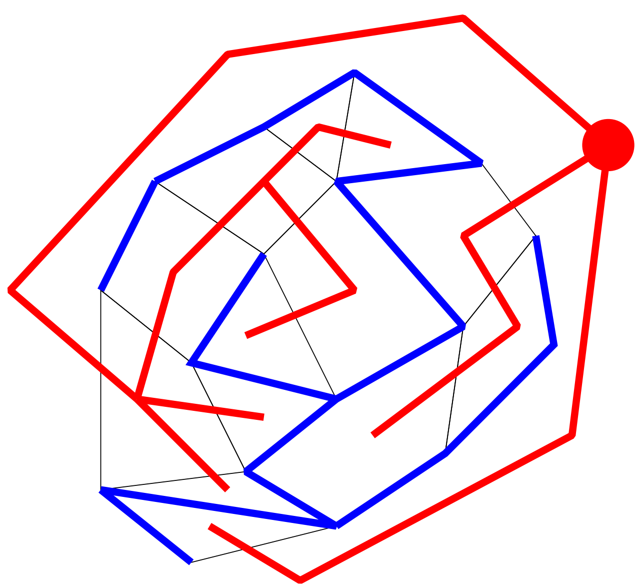

Figure 2: the trees T and T^*.

2 Kastelyn's Theorem

Let G be a nice graph with a Pfaffian orientation. Assume that G has an even number of vertices. Let $v_1, ..., v_{2n}$ be the vertices of G. Let A denote the signed adjacency matrix of G. This means that

- $A_{ij} = 0$ if no edge joins v_i to v_j.
- $A_{ij} = 1$ if the edge joining v_i to v_j is oriented from v_i to v_j.
- $A_{ij} = -1$ if the edge joining v_i to v_j is oriented from v_j to v_i.

A perfect matching of G is a collection of edges $e_1, ..., e_n$ of G such that every vertex belongs to exactly one e_j. Let $P(G)$ denote the number of distinct perfect matchings of G. These notes are devoted to proving:

Theorem 2.1 (Kastelyn) $P(G) = \sqrt{|\det(A)|}$.

Let A^t denote the transpose of a matrix A. The matrix A is called *skew-symmetric* if $A^t = -A$. This is equivalent to the condition $A_{ij} = -A_{ji}$ for all i, j. By construction, the signed adjacency matrix A is skew-symmetric.

Let A be a $2n \times 2n$ skew symmetric matrix. The *Pfaffian* of A is defined as the sum

$$p(A) = \frac{1}{2^n n!} \sum_\sigma \text{sign}(\sigma) \prod_{i=1}^n A_{\sigma(2i-1),\sigma(2i)}. \tag{1}$$

The sum is taken over all permutations of the set $\{1, ..., 2n\}$.

Below, we'll prove the following result, known as *Muir's Identity*.

$$(p(A))^2 = \det(A). \tag{2}$$

Given Muir's Identity, Kastelyn's Theorem can be reformulated as follows.

Theorem 2.2 *Let G be a nice graph, equipped with any Pfaffian orientation. Let A be the corresponding signed adjacency matrix. Then $P(G) = |p(A)|$.*

First I'll prove Theorem 2.2 and then I'll derive Muir's Identity.

3 Proof Theorem 2.2

Given that A is a skew-symmetric matrix, there is some redundancy in the definition of the Pfaffian. We write our permutations as

$$(1, ..., 2n) \to (i_1, j_1, i_2, j_2, ..., i_n, j_n). \tag{3}$$

Say that two permutations are *equivalent* if one can get from one to the other by permuting the (i, j) pairs and/or switching elements within the pairs. Each equivalence class of permutations has $2^n n!$ members. Moreover, each equivalence class has a unique member such that $i_1 < ... < i_n$ and $i_k < j_k$ for all k. Let $\mathcal{S}$ denote the set of these special permutations.

Two equivalent permutations contribute the same term in $p(A)$, thanks to the skew symmetry of A. Therefore, we have the alternate definition

$$p(A) = \sum_{\sigma \in \mathcal{S}} \text{sign}(\sigma) \prod_{i=1}^{n} A_{\sigma(2i-1),\sigma(2i)}. \tag{4}$$

The *terms* in Equation 4 are in one-to-one correspondence with the perfect matchings. To prove Theorem 2.2, we just have to prove that all such terms have the same sign. Let $s(M)$ be the term of Equation 4 corresponding to the perfect matching M. We orient the edges of M according to the Pfaffian orientation. Call an edge of M *bad* if it points from a higher indexed vertex to a lower indexed vertex. The sign of $s(M)$ just depends on the parity of the number of bad edges. We just have to prove that every two perfect matchings M_1 and M_2 have the same parity of bad edges.

The symmetric difference $M_1 \Delta M_2$ is a finite union of (alternating) cycles of even length. We can produce a new perfect matching M' by switching the matching on one or more of the cycles of $M_1 \Delta M_2$. Thus, M_1 and M_2 can be "joined" by a chain of perfect matchings, each of which differ by a single even alternating cycle. Hence, it suffices to consider the case when $M_1 \Delta M_2$ is a single cycle C.

If we permute the indices of G and recompute the signs, all permutations change by the same global sign. So, it suffices to consider the case when the vertices of C are given as $1, ..., 2k$ in counterclockwise order. The edges of $M_1|_C$ are (12), (34), etc. so the bad edges of $M_1|_C$ are precisely the clockwise edges. If $M_1|_C$ has an odd (even) number of bad edges then $M_1|_C$ has an odd (even) number of clockwise edges. The edges of $M_2|_C$ are (23), (45), ..., (k1). Hence, if $M_2|_C$ has an odd (even) number of bad edges then $M_2|_C$ has an even (odd) number of clockwise edges.

The cycle C surrounds some edges common to M_1 and M_2. Therefore C surrounds an even number of vertices. Below we will prove that C must have an odd number of clockwise edges. Hence, the number of clockwise edges in the two cases has opposite parity. Hence, the number of bad edges in the two cases has the same parity. Hence $s(M_1) = s(M_2)$.

The following Lemma completes the proof of Theorem 2.2.

Lemma 3.1 (Kastelyn) *Suppose that C is a cycle that surrounds an even number of vertices. Then C has an odd number of clockwise edges.*

Proof: Let $n(C)$ denote the number of clockwise edges of C plus the number of vertices contained in the interior of the region bounded by C. It suffices to prove that $n(C)$ is always odd.

Every nice planar graph G' can be extended to a triangulated disk G, and every Pfaffian orientation on G' extends to a Pfaffian orientation on G. (When each new edge is drawn, breaking up an elementary cycle, there is a unique way to orient the edge to keep the condition on the two smaller cycles that are created.) So, it suffices to prove this result when C is the boundary of a triangulated disk Δ.

There are two cases to consider, Suppose Δ has an interior edge e which joins two vertices of C. Then we can write $\Delta = \Delta_1 \cup \Delta_2$, where Δ_1 and Δ_2 are two smaller disks joined along e. Let C_j be the boundary of Δ_j. Here we have $n(C) - 1 = n(C_1) + n(C_2)$ because e is counted exactly once on the right hand side. In this case, we see by induction that $n(C)$ is odd.

The remaining case to consider is when Δ has a triangle T having an edge on C and one interior vertex, as shown in Figure 3. In this case, let $\Delta' = \Delta - T$. Let C' be the boundary of Δ'. A short case by case analysis shows that $N(C) = N(C')$ or $N(C) = N(C') - 2$. The result again holds by induction. ♠

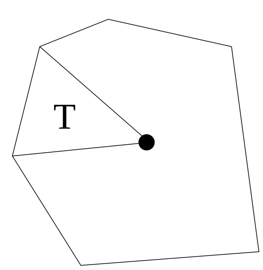

Figure 3: Chopping off a triangle

4 Muir's Identity

The remainder of these notes is devoted to proving Muir's Identity.

Call the matrix A *special* if it has the form

$$A = \begin{bmatrix} M_1 & 0 & \dots & 0 & 0 & 0 \\ 0 & M_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & M_3 & 0 & \dots & 0 \\ \dots & & & & & \end{bmatrix}, \qquad M_k = \begin{bmatrix} 0 & a_k \\ -a_k & 0 \end{bmatrix} \tag{5}$$

From Equation 4 we have $p(A) = a_1 ... a_n$. Also $\det(A) = (a_1 ... a_n)^2$. So, Muir's identity holds when A is a special matrix.

Below we will prove

Lemma 4.1 *Let $\mathcal{X}$ denote the set of all $2n \times 2n$ skew symmetric matrices. There is an open $\mathcal{S} \subset \mathcal{X}$ with the following property. If $A \in \mathcal{S}$ then there is a rotation B (depending on A) such that BAB^{-1} is a special matrix.*

After we prove Lemma 4.1, we will establish the following transformation law. If B is any $2n \times 2n$ matrix, then

$$p(BAB^t) = \det(B) p(A). \tag{6}$$

Let $A \in \mathcal{S}$, the set from Lemma 4.1. We have

$$(p(A))^2 =_1 (p(BAB^{-1}))^2 =_2 \det(BAB^{-1}) =_3 \det(A).$$

Equality 1 comes from the Equation 6 and from the fact that $B^t = B^{-1}$ when B is a rotation. Equality 2 is Muir's identity for special matrices. Equality 3 is a familiar fact about determinants. Since Muir's identity is a polynomial identity, and holds on an open subset of skew-symmetric matrices, the identity holds for all skew-symmetric matrices.

5 Proof of Lemma 4.1

The best way to do this is to work over the complex numbers, and then drop back down to the reals at the end.

Let $\langle , \rangle$ be the standard Hermitian form on $\boldsymbol{C}^{2n}$. That is

$$\langle (z_1, ..., z_{2n}), (w_1, ..., w_{2n}) \rangle = \sum_{i=1}^{2n} z_i \overline{w}_i. \tag{7}$$

Here, the bar denotes complex conjugation. Here are some basic properties of this object.

- When $v, w \in \boldsymbol{R}^{2n}$, we have $\langle v, w\rangle = v \cdot w$. So, the Hermitian form is kind of an enhancement of the dot product.
- $\langle w, v\rangle = \overline{\langle v, w\rangle}$.
- $\langle av_1 + v_2, w\rangle = a\langle v_1, w\rangle + \langle v_2, w\rangle$.
- Let M^* denote the conjugate transpose of M. Then $\langle Mv, w\rangle = \langle v, M^*w\rangle$.

These properties will be used in the next lemma.

Lemma 5.1 *If A is real and skew symmetric (so that $A^* = -A$) then the eigenvalues of A are pure imaginary.*

Proof: Let λ be an eigenvalue and v a corresponding eigenvector. We have

$$\lambda\langle v, v\rangle = \langle Av, v\rangle = \langle v, A^*v\rangle = \langle v, -\lambda v\rangle = -\overline{\lambda}\langle v, v\rangle. \tag{8}$$

Hence $\overline{\lambda} = -\lambda$. This is only possible if λ is pure imaginary. ♠

Let $\mathcal{S}$ denote the set of real skew-symmetric matrices having eigenvalues of the form $\lambda_1, \overline{\lambda}_1, ..., \lambda_n, \overline{\lambda}_n$, where $0 < |\lambda_1| < ... < |\lambda_n|$. Note that $\mathcal{S}$ is nonempty, because one can easily construct a special matrix in $\mathcal{S}$. Note also that $\mathcal{S}$ is open, because the eigenvalues vary continuously. So, $\mathcal{S}$ contains a nonempty open set.

Let A be any element of $\mathcal{S}$. Let $E_k \subset \boldsymbol{C}^{2n}$ denote the subspace of vectors of the form $av + bw$ where $A(v) = \lambda_k(v)$ and $B(w) = \overline{\lambda}_k(w)$. By construction, E_k is a 2-dimensional subspace over $\boldsymbol{C}$.

Lemma 5.2 *Suppose $i \neq j$. If $v \in E_i$ and $w \in E_j$ then $\langle v, w\rangle = 0$.*

Proof: It suffices to prove that $\langle v_1, v_2\rangle = 0$ when v_1 and v_2 are eigenvalues of A corresponding to eigenvalues having different magnitudes. We have

$$\lambda_1\langle v_1, v_2\rangle = \langle A(v_1), v_2\rangle = \langle v_1, A^*(v_2)\rangle = -\langle v_1, A(v_2)\rangle = -\overline{\lambda}_2\langle v_1, v_2\rangle.$$

Hence, either $|\lambda_1| = |\lambda_2|$ or $\langle v_1, v_2\rangle = 0$. ♠

Let $V_k = E_k \cap \boldsymbol{R}^{2n}$.

Lemma 5.3 *V_k is 2-dimensional as a real vector space.*

Proof: Note that E_k is invariant under complex conjugation. V_k is at least 2 dimensional because it contains all vectors of the form $v+\overline{v}$ with $v \in E_k$. On the other hand V_k cannot be more than 2-dimensional because the $\boldsymbol{C}$-span of V_k is contained in E_k. ♠

For real vectors v, w, recall that $\langle v, w\rangle$ is just the dot product. Hence $v \cdot w = 0$ when $v \in V_i$ and $w \in V_j$. So the n planes $V_1, ..., V_n$ are pairwise perpendicular. Note also that $A(V_k) = V_k$. So, we have succeeded in finding n pairwise perpendicular and A-invariant 2-planes in $\boldsymbol{R}^{2n}$.

Lemma 5.4 *A acts on each V_k as a dilation by $|\lambda_k|$ followed by a 90 degree rotation.*

Proof: Note that A^2 has all negative real eigenvalues. On E_k, we have $A^2(v) = -|\lambda_k|^2(v)$. The lemma follows immediately from this. ♠

Now for the end of the proof: Given the properties of $V_1, ..., V_n$, we can find a rotation B such that

$$B(V_k) = W_k := \mathrm{span}(e_{2k-1}, e_{2k}), \qquad k = 1, ..., n. \tag{9}$$

The map BAB^{-1} preserves W_k and acts on W_k as the composition of a dilation and a 90 degree rotation. From here, it is easy to see that BAB^{-1} must be a special matrix.

6 The Transformation Law

6.1 Brute Force Approach

By an obvious scaling argument, it suffices to establis Step 1 when $\det(B) = 1$. As is well known from Gaussian elimination, every such B is the product of elementary matrices. Using the fact that $\det(E_1E_2) = \det(E_1)\det(E_2)$, it suffices to establish Step 1 for elementary matricess. This is a fairly easy calculation which you might prefer to the more abstract argument given below.

6.2 The Main Argument

Let $V = \boldsymbol{R}^{2n}$ and let V^* denote the dual space. Let $e^1, ..., e^{2n}$ denote the standard basis for V^*.

Given the skew-symmetric matrix A, introduce the 2-form

$$[A] = \sum_{i<j} A_{ij}\ e^i \wedge e^j. \tag{10}$$

Let $\wedge^n[A] = [A] \wedge ... \wedge [A]$, the n-fold wedge product. $\wedge^n[A]$ is an alternating $2n$-form, and hence is some multiple of $e^1 \wedge ... \wedge e^{2n}$.

Lemma 6.1 $\wedge^n[A] = n!P(A)e^1 \wedge ... \wedge e^{2n}$.

Proof: Say that the permutation in Equation 3 is *semi-special* if $i_k < j_k$ for all k. Let $\mathcal{SS}$ denote the set of semi-special permutations. Every permutation is equivalent to a semi-special permutation through the operation of permuting the (i,j) pairs. Hence

$$n!p(A) = \sum_{\sigma\in\mathcal{SS}} \text{sign}(\sigma) \prod_{i=1}^{n} A_{\sigma(2i-1),\sigma(i)}. \tag{11}$$

But the right hand side is exactly what you get when you expand out $\wedge^n[A]$. ♠

Let $B : \boldsymbol{R}^{2n} \to \boldsymbol{R}^{2n}$ be a linear map, represented by a square matrix. Let $B^*[A]$ denote the pull-back of $[A]$ by B. We compute

$$B^*([A])(e_a, e_b) := [A](B(e_a), B(e_b)) = \sum_{i<j} A_{ij}\ e^i \wedge e^j(B(e_a), B(e_b)) =$$

$$\sum_{i<j} A_{ij}\ e^i \wedge e^j\Big(\sum_k B_{ak}e_k, \sum_l B_{bj}e_l\Big) = \sum_{i<j} A_{ij}B_{ai}B_{jl} = (BAB^t)_{ab}. \tag{12}$$

In short, $B^*[A] = [BAB^t]$. Therefore,

$$\det(B) \wedge^n [A] =^1 B^*(\wedge^n[A]) =^2 \wedge^n(B^*([A])) = \wedge^n[BAB^t]. \tag{13}$$

Equality 1 is the transformation law for $2n$-forms. Equality 2 is the compatibility of the wedge product with pullbacks.

In short, $\wedge^n[BAB^t] = \det(B) \wedge^n [A]$. Equation 6 follows from this result and from Lemma 6.1.

6.3 Forms

Now I'll give details on tensors which should explain the argument above.

Let $V = \boldsymbol{R}^{2n}$. Let V^* denote the vector space of linear functions from V to $\boldsymbol{R}$. The space V^* is called *the dual* of V. A canonical basis for V^* is given by $e^1, ..., e^{2n}$, where e^j is the function such that

$$e^j(a_1, ..., a_{2n}) = a_j. \tag{14}$$

A *k-tensor* is a map $\phi : V \times ... \times V \to \boldsymbol{R}$ such that

$$\phi(*, ..., *, av+w, *, ..., *) = a\phi(*, ..., *, v, *, ..., *)+\phi(*, ..., *, w, *, ..., *). \tag{15}$$

In other words, if you hold all positions fixed but one, then ϕ is a linear function of the remaining variable. The product $V \times ... \times V$ is supposed to be the k-fold product.

The k-tensor ϕ is said to be a *k-form* if

$$\phi(*, ..., *, a_i, *, ..., *, a_j, *, ..., *) = -\phi(*, ..., *, a_j, *, ..., *, a_i, *, ..., *). \tag{16}$$

In other words, if you switch two of the entries, the sign changes.

If ϕ_1 and ϕ_2 are k-forms, then so are $r\phi_1$ and $\phi_1 + \phi_2$. In other words, the set of all k-forms makes a vector space. This vector space is denoted $\wedge^k(V)$. A nice element of $\wedge^k(V)$ is given by

$$\omega = e^{i_1} \wedge ... \wedge e^{i_k}, \qquad i_1 < ... < i_k. \tag{17}$$

This tensor has the following description. If we want to evaluate $\omega(v_1, ..., v_k)$, we make a $2n \times k$ matrix having rows $v_1, ..., v_k$. Then we take the square minor made from columns $i_1, ..., i_k$, then we take the determinant of this minor. Familiar properties of the determinant guarantee that ω really is a k-form. Call ω an *elementary* k-form.

Lemma 6.2 *The elementary k-forms make a basis for $\wedge^k(V)$.*

Proof: Any k-form is determined by what it does to the k-tuples $(e_{i_1}, ..., e_{i_k})$, as the indices range over all possibilities. Hence, the elementary forms span $\wedge^k(V)$. Also, the form $e^{i_1} \wedge ... \wedge e^{i_k}$ assigns 1 to $(e_{i_1}, ..., e_{i_k})$ and 0 to all other such lists. Hence this particular elementary k-form is not a linear combination of the others. But this means that the elementary k-forms are linearly independent. ♠

6.4 The Wedge Product

Now we know that $\wedge^k(V)$ has dimension $2n$ choose k. In particular, $\wedge^{2n}(V)$ has dimension 1: all $2n$-forms are multiples of $e^1 \wedge ... \wedge e^{2n}$.

Given two lists $I = \{i_1 < ... < i_a\}$ and $J = \{j_1 < ... < j_b\}$, let K denote the sorted list of $I \cup J$. We define

- $\sigma(I, J) = 0$ if $I \cap J \neq \emptyset$.
- $\sigma(I, J) = 1$ if an even permutation sorts the elements $i_1, ..., i_a, j_1, ..., j_b$.
- $\sigma(I, J) = -1$ if an odd permutation sorts the elements $i_1, ..., i_a, j_1, ..., j_b$.

There is a bi-linear map

$$\wedge : \wedge^a(V) \times \wedge^b(V) \tag{18}$$

defined as follows.

$$(e^{i_1} \wedge ... \wedge e^{i_a}) \wedge (e^{j_1} \wedge ... \wedge e^{j_b}) = \sigma(I, J) e^{k_1} \wedge ... \wedge e^{k_{a+b}}. \tag{19}$$

The map is extended to all of $\wedge^a(V) \times \wedge^b(V)$ using bi-linearity. This map is known as the *wedge product*. The wedge product is associative, so that $(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma)$.

Another formula for the wedge product is

$$\alpha \wedge \beta(v_1, ..., v_a, w_{a+1}, ..., w_{a+b}) =$$

$$\frac{1}{(a+b)!} \sum_{\sigma} \operatorname{sign}(\sigma) \alpha(v_{\sigma(1)}, ..., v_{\sigma(a)}) \times \beta(w_{\sigma(a_1)}, ..., w_{\sigma(a+b)}) \tag{20}$$

The sum is over all permutations of the set $1, ..., a+b$. The $(\times)$ in this equation is just ordinary multiplication. One can check, using properties of determinants, that this formula holds when α and β are elementary forms. Then, the general case follows from the linearity of Equation 20.

As in the short version of Step 1, we define $[A] \in \wedge^2(V)$ to be the 2-form

$$[A] = \sum_{i<j} A_{ij} e^i \wedge e^j. \tag{21}$$

Note that $[A] \wedge ... \wedge [A]$ (n-times) is an element of $\wedge^{2n}(V)$. Hence, there is a constant C such that $[A] \wedge ... \wedge [A] = Ce^1 \wedge ... \wedge e^{2n}$. In Lemma 6.1 we just expand this out, using the multi-linearity of the wedge product, to find the constant C.

6.5 Transformation Laws

Nothing we say here has to do with the dimension being even, so we set $m = 2n$ and remark that everything works for m odd as well. Suppose that T is a linear transformation and ω is a k-form. We have the general definition

$$T^*(\omega)(v_1, ..., v_k) = \omega(T(v_1), ..., T(v_k)). \tag{22}$$

As a special example, we have

$$T^*(e^1 \wedge ... \wedge e^m)(e_1, ..., e_m) = e^1 \wedge ... \wedge e^m(T(e_1, ..., t(2_m)) =$$

$$e^1 \wedge ... \wedge e^m\Big(\sum T_{1j}e_j, ..., \sum T_{nj}e_j\Big) =$$

$$\Big(\sum_\sigma \text{sign}(\sigma) \prod T_{i,\sigma(i)}\Big)e^1 \wedge ... \wedge e^m = \det(T)e^1 \wedge ... \wedge e^m.$$

The sum is taken over all permutations. Since $\wedge^m(V)$ is 1-dimensional, and the pull back operation respects scaling, we have the simpler formula

$$T^*(\omega) = \det(T)\omega, \tag{23}$$

for any m-form. This explains Equality 1 in Equation 13.

It is an easy consequence of Equation 20 that

$$T^*(\alpha \wedge \beta) = T^*(\alpha) \wedge T^*(\beta). \tag{24}$$

From the associativity of the wedge product, we get

$$T^*(\alpha_1 \wedge ... \wedge \alpha_k) = T^*(\alpha_1) \wedge ... \wedge T^*(\alpha_k). \tag{25}$$

This explains Equality 2 in Equation 13.