A combinatorial trace formula*

F. R. K. Chung
University of Pennsylvania
Philadelphia, Pennsylvania 19104

S.-T. Yau
Harvard University
Cambridge, Massachusetts 02138

Abstract

We consider the trace formula in connection with the Laplacian of a graph. This can be viewed as the combinatorial analog of Selberg's trace formula, which has influenced many areas of number theory, representation theory and geometry. We start with a general graph with any specified set of edges. Therefore the trace formula we consider is in full generality. In particular, we consider lattice graphs, their subgraphs and higher-dimensional generalizations. By using the analog of the Poisson inversion formula, several hypergeometric equalities are derived.

1 Introduction

We consider a graph $G = (V, E)$ with vertex set $V = V(G)$ and edge set $E = E(G)$. We assume G contains no loops or multiple edges (the definitions can be easily generalized to weighted graphs with loops, cf. [5]). We will mainly consider finite graphs (while an infinite graph is viewed as taking the "limit" of a class of finite graphs). We define the matrix L with rows and columns indexed by vertices of G as follows.

$$L(u,v) = \begin{cases} d_v & \text{if } u = v \\ -1 & \text{if } u \text{ and } v \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

where d_v denotes the degree of v.

Let T denote the diagonal matrix with the (v, v)-entry having value d_v. The Laplacian $\mathcal{L}$ of G is defined to be

$$\mathcal{L}(u,v) = \begin{cases} 1 & \text{if } u = v \text{ and } d_v \neq 0 \\ -\dfrac{1}{\sqrt{d_u d_v}} & \text{if } u \text{ and } v \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

*This paper has appeared in *Tsing Hua Lectures on Geometry and Analysis*, (ed. S.-T. Yau), International Press, Cambridge, Massachusetts, 1997, 107-116.

When G is k-regular (i.e., $d_v = k$ for all v), it is easy to see that

$$\mathcal{L} = I - \frac{1}{k}A$$

where A is the adjacency matrix of G. In general, the eigenvalues of $\mathcal{L}$ and L can be quite different.

The eigenvalues of $\mathcal{L}$ are denoted by $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-1}$. If G is connected, we have $0 < \lambda_1$. Various properties of the λ_1's can be found in [3] .

For a graph G, we consider the heat kernel H_t, which is defined for $t \geq 0$ as follows:

$$\begin{aligned} H_t &= \sum_i e^{-\lambda_i t} P_i \\ &= e^{-t\mathcal{L}} \\ &= I - t\mathcal{L} + \frac{t^2}{2}\mathcal{L}^2 - \ldots \end{aligned} \tag{1}$$

where P_i denotes the projection into the eigenspace associated with eigenvalue λ_i. In particular,

$$H_0 = I.$$

and H_t satisfies the heat equation

$$\frac{\partial H_t}{\partial t} = -\mathcal{L}H_t$$

Here we state some useful facts about the heat kernel.

Fact 1: For any two vertices x and y, we have

$$H_t(x, y) \geq 0$$

Fact 2: For any two vertices x and y, we have

$$H_t(x, y) = \sum_j e^{-t\lambda_j} \phi_j(x)\phi_j(y)$$

where ϕ_k denotes the eigenfunction corresponding the the eigenvalue λ_k of the Laplacian $\mathcal{L}$ and ϕ_k's are orthonormal.

We omit the proofs for both facts since it can be found in [6]. Fact 2 looks hard but it can be easily proved using the definitions while Fact 1 looks easy but its proof is quite involved.

We remark that we can consider induced subgraphs of a graph and heat kernels with the Dirichlet boundary condition or Neumann boundary conditions [5] . However, for simplicity of exposition, we will deal with graphs with no boundaries in this paper.

2 The trace formula

For a graph G with heat kernel H_t and eigenvalues λ_i, the simplest version of the trace formula is:

$$Tr\ (H_t) = \sum_x H_t(x,x) = \sum_j e^{-\lambda_j t} \tag{2}$$

In general, the heat kernel is not easy to determine. However, the trace can be computed in polynomial time (of the same order of complexity as matrix multiplication). We remark that a good part of spectral geometry is in estimating the heat kernel from which we can extract various invariants of the Riemannian manifold. In a way, the trace formula captures the essential part of the heat kernel and it can often serve the same purpose. The combinatorial trace formula provides an effective tool to capture the major invariants and properties of a graph just as the Selberg's trace formula does for algebraic structures and geometric surfaces.

If the graph is vertex transitive, (i. e. for any two verticies u and v, there is an edge-preserving automorphism mapping u to v.) then for all v, $H_t(x,x)$ has the same value and thus is easier to evaluate.

Perhaps the simplest vertex-transitive graph is an n-cycle. As it turns out, the heat kernel for an n-cycle can be deduced from the heat kernel of an infinite path. In fact, the heat kernel $\mathbf{H}(x,x)$ of an infinite paths is fundamental in the sense that it can be used to determine the heat kernels of numerous graphs such as n-cycles, n-paths P_n, higher dimensional lattice graphs and some of their induced subgraphs. The importance of the heat kernel of an infinite path is perhaps due to the fact that it is the discrete analog of a line which the higher dimensional Riemannian manifolds are based upon.

We will first derive the heat kernel for an infinite path. Based on it, we will derive trace formula for cycles, paths, lattice graphs and "quotient" graphs which we will describe in Section 6.

Before we proceed, we state a simple but useful combinatorial observation.

Lemma 1 *In a k-regular graph G, suppose $W_s = (v_0, \cdots, v_s)$ is a walk of length s from vertex v_0, through vertices v_j to v_s such that $v_j = v_{j+1}$ or $\{v_j, v_{j+1}\}$ is an edge. Then the heat kernel $H_t(x,y)$ of G satisfies*

$$H_t(x,y) = \sum_s \frac{t^s}{s!}\Big(\sum_r \Big(\frac{-1}{k}\Big)^r \rho(s,r)\Big)$$

where $\rho(s,r)$ denotes the number of walks of length s from $x = v_0$ to $y = v_s$ containing exactly r edges.

Proof: This follows from the fact that

$$H = I - t\mathcal{L} + \frac{t^2}{2}\mathcal{L}^2 + \cdots$$

and $\mathcal{L}^s(x, y)$ is exactly the sum of weights over all walks of length s from x to y where the weight of a path of length s containing exactly r edges is just

$$\prod_{j=0}^{s-1} \mathcal{L}(v_j, v_{j+1}) = (\frac{-1}{k})^r.$$

3 The heat kernel for the one-dimensional lattice graphs

First we consider the one-dimensional case. The heat kernel for an infinite path can be viewed as taking the limit of an n-cycle, C_n. We use the notation that vertices are labeled by integers and x is adjacent to $x + 1$ and $x - 1$.

Theorem 1 *In an infinite path, the heat kernel satisfies:*

$$\mathbf{H}_t(x, x) = \sum_{k \geq 0} \frac{\binom{2k}{k}}{k!} (\frac{-t}{2})^k \tag{3}$$

Proof: We use Lemma 1 and also we use the fact that $\mathcal{L} = I - Y/2 - Y^{-1}/2$ where Y is the cyclic operator $Y(x) = x + 1$. The coefficient of t^k in the expansion of $\mathbf{H}_t(x, x)$ in (1) is exactly $1/k!$ multiplied by the coefficient of Y^0 in $(I - Y/2 - Y^{-1}/2)^k$. By collecting the coefficients term by term, Theorem 1 is proved.

From Theorem 1, we can see that $\mathbf{H}_t(x, x)$ is exactly a confluent hypergeometric series ${}_1F_1(1/2; 1; -2t)$ and is closely related to the modified Bessel functions $I_\nu(z)$. That is

$$\mathbf{H}_t(x, x) = F(1/2; 1; -2t) = e^{-t} I_0(-t) \tag{4}$$

It is known that the above confluent hypergeometric series and the modified Bessel function do not have a closed form formula. The usual asymptotical expansions for I_v or for the confluent hypergeometric series (e.g. [5]) do not seem to have the following useful forms as stated in the following two theorems:

Theorem 2 *In an infinite path, the heat kernel satisfies:*

$$\mathbf{H}_t(x, x) = \frac{\sqrt{2}}{\pi\sqrt{t}} \int_0^{\sqrt{2t}} \frac{e^{-y^2}}{\sqrt{1 - \frac{y^2}{2t}}} dy$$

Proof: For an n-cycle, the eigenvalues of the Laplacian are

$$\lambda_k = 1 - \cos\frac{2\pi k}{n} = 2\sin^2\frac{\pi k}{n}$$

for $k = 0, 1, \cdots, n-1$. Therefore we have

$$\begin{aligned}
\mathbf{H}_t(x,x) &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} e^{-t\lambda_k} \\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} e^{-t(1-\cos\frac{2\pi k}{n})} \\
&= \frac{1}{\pi}\int_0^{\pi} e^{-2t\sin^2 x}dx \\
&= \frac{2}{\pi}\int_0^{\pi/2} e^{-2t\sin^2 x}dx \\
&= \frac{\sqrt{2}}{\pi\sqrt{t}}\int_0^{\sqrt{2t}} \frac{e^{-y^2}}{\sqrt{1-\frac{y^2}{2t}}}dy
\end{aligned}$$

For t large and a fixed, we can use the asymptotic expansion of $I_0(-t)$:

$$I_0(-t) \quad \sim \quad \frac{e^t}{\sqrt{2\pi t}}\left(1 + \sum_{m\geq 1}\frac{1^2\cdot 3^2\cdots(2m-1)^2}{(-2t)^m}\right)$$

Therefore we have

Theorem 3 *When t is sufficiently large, the heat kernel of an infinite path satisfies*

$$\mathbf{H}_t(x,x) = \frac{1}{\sqrt{2t\pi}}(1 + O(\frac{1}{t}))$$

Next, we consider the off-diagonal terms of the heat kernel of the infinite path, that is $\mathbf{H}_t(x,y)$.

Theorem 4 *In an infinite path, the hear kernel satisfies, for any integer $a \geq 0$,*

$$\mathbf{H}_t(x,x+a) = \mathbf{H}_t(x,x-a) = (-1)^a\sum_{k\geq a}\frac{\binom{2k}{k+a}}{k!}(\frac{-t}{2})^k$$

Proof: The proof is quite similar to that of Theorem 1. We here focus on the coeffient of Y^a in $\mathcal{L}^k = (I - Y/2 - Y^{-1}/2)^k$ for each k.

We note that $\binom{m}{n} = 0$ for integers $n > m$. So the sum in the statement of Theorem 4 can be taken over all non-negative integers k. Using Theorem 1, we here state the relation of $\mathbf{H}_t(x, x+a)$ to a confluent hypergeometric series and the modified Bessel functions $I_\nu(z)$.

Theorem 5

$$\mathbf{H}_t(x, x+a) = (t/2)^a/a!\ {}_1F_1(1/2+a; 1+2a; -2t) = (-1)^a e^{-t} I_a(-t)$$

Again the above hypergeometric series in Thereom 4 has no closed formula. The integral similar to that in Theorem 2 is somewhat complicated. However, the asymptotical estimates for large $t >> a^2$ have a very simple and elegant form:

Theorem 6 *The heat kernel of an infinite path satisfies*

$$\mathbf{H}_t(x, x+a) = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} e^{-2t\sin^2 x}\cos 2ax dx$$

Proof: We follow the proof of Theorem 2. In an n-cycle, the eigenfunction ϕ_k associated with the eigenvalue $\lambda_k = 1 - \cos\frac{2\pi k}{n}$ for $k = 0, 1, \cdots, n-1$ is $\phi_k(j) = \exp(\frac{2\pi ikj}{n})/\sqrt{n}$. We have

$$\begin{aligned}
\mathbf{H}_t(x, x+a) &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} e^{-t\lambda_k} e^{\frac{2\pi ika}{n}} \\
&= \frac{1}{\pi}\int_0^{\pi} e^{-2t\sin^2 x + 2iax} dx \\
&= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} e^{-2t\sin^2 x}\cos 2ax dx
\end{aligned}$$

Theorem 7 *For a fixed a and $t >> a^2$, we have*

$$\mathbf{H}_t(x, x+a) = \frac{e^{-\frac{a^2}{2t}}}{\sqrt{2\pi t}}(1 + O(\frac{a}{\sqrt{t}}))$$

Proof: Under the assumption that a is fixed and $t >> a^2$, we use the known estimates on Bessel functions $I_\nu(z)$ and $J_\nu(z)$ [8]:

$$\begin{aligned}
I_a(-t) &= (-i)^a J_a(-it), \\
J_\nu(\nu\ \mathrm{sech}\ \alpha) &= \frac{e^{\nu(\tanh\alpha - \alpha)}}{\sqrt{2\pi\nu\tanh\alpha}}(1 + O(\frac{1}{\sqrt{\nu\tanh\alpha}})).
\end{aligned}$$

We choose

$$\begin{aligned}
a\ \mathrm{sech}\ \alpha &= -i\ t, \\
\tanh\alpha &= \frac{\sqrt{a^2+t^2}}{a}, \\
-\alpha &= \frac{3\pi i}{2} + \ln\frac{\sqrt{a^2+t^2} - a}{t}
\end{aligned}$$

By substituting into $\mathbf{H}_t(x, x+a)$, we get

$$\mathbf{H}_t(x, x+a) \quad = \quad \frac{e^{-\frac{a^2}{2t}}}{\sqrt{2\pi t}}(1 + O(\frac{a}{\sqrt{t}}))$$

Using Theorem 4 and 6, we deduce the heat kernel $H_t(x, y, C_n)$for an n-cycle:

Theorem 8 *In the n-cycle, the heat kernel satisfies*

$$H_t(x, y, C_n) = \sum_{k=-\infty}^{\infty} \mathbf{H}_t(x, y + kn)$$

Proof: Using the observation in Lemma 1, we note that all walks from x to x in the cycle C_n can be identified with walks in the infinite paths from x to $x + nk$, for some k. The sum of weights from walks from x to $x + nk$ contribute to $\mathbf{H}_t(x, x + kn)$. As an immediate consequence of Theorem 4 and 8, we have the following trace formula for an n-cycle:

Theorem 9

$$\sum_{k=0}^{\infty} \sum_{j=-\infty}^{\infty} (-1)^{nj} \frac{\binom{2k}{k+nj}(-\frac{t}{2})^k}{k!} = \frac{1}{n} \sum_{k=0}^{n-1} e^{-2t \sin^2 \frac{\pi k}{n}}$$

4 The heat kernel for $k-$dimensional lattice graphs

We define the lattice graph $C_n^{(k)}$ to be the cartesian product of k copies of an n-cycle. The infinite lattice graph $P^{(k)}$ is just by taking the limit of $C_n^{(k)}$ as n approaches infinity. We first consider the $2-$dimensional cases. For $P^{(2)}$, each vertex is labelled by $(x, y), x, y \in \mathbb{Z}$. (x, y) is adjacent to $(x+1, y), (x-1, y), (x, y+1)$ and $(x, y-1)$.

Theorem 10 *In an infinite $2-$dimensional lattice graph, the heat kernel $\mathbf{H}_t^{(2)}$ satisfies*

$$\begin{aligned} \mathbf{H}_t^{(2)}((x, y), (x+a, y+b)) \quad &= \quad (-1)^{a+b} \sum_k \sum_j \binom{k}{j} \binom{2j}{j+a} \binom{2(k-j)}{k-j+b} \frac{(-t/4)^k}{k!} \\ &= \quad \mathbf{H}_{t/2}(x, x+a) \mathbf{H}_{t/2}(y, y+b) \end{aligned}$$

where $\mathbf{H}_t$ is the heat kernel for an infinite path.

Proof: The proof follows from the fact that

$$\mathcal{L} = I - \frac{1}{4}Y_1 - \frac{1}{4}Y_1^{-1} - \frac{1}{4}Y_2 - \frac{1}{4}Y_2^{-1}$$

where

$$Y_1(x,y) = (x+1,y),\ Y_2(x,y) = (x,y+1)$$

Using the expansion of $e^{-t\mathcal{L}}$, and Theorem 1, we see that $\mathbf{H}_t^{(2)}((x,y),(x+a,y+b))$ can be factored into $\mathbf{H}_{t/2}(x,x+a)$ and $\mathbf{H}_{t/2}(y,y+b)$.

Therefore we have

Theorem 11 *For two vertices u and v in a $2-$dimensional lattice graph, we have*

$$\mathbf{H}_t^{(2)}(u,v) = \frac{e^{-\frac{\|u-v\|^2}{t}}}{t\pi}(1 + O(\frac{\|u-v\|}{\sqrt{t}}))$$

where $\|\cdot\|$ denotes the L_2-norm.

Proof: By choosing $u = (x,y)$ and $v = (x+a, y+b)$, Theorem 11 is a consequence of Theorem 10 and Theorem 5.

In a similar way, we have

Theorem 12 *The heat kernel for P^k satisfies, for $a = (a_1, \dots, a_k)$, $a_i \in \mathbb{Z}$,*

$$K_t(a) = K_{\frac{t}{k}}(a_1)\dots K_{\frac{t}{k}}(a_k) \tag{5}$$

where $K_t(a) = \mathbf{H}_t^{(k)}(x, x+a)$.

Similar to the proof of Theorem 11, we have

Theorem 13

$$\mathbf{H}_t^{(k)}(u,v) = \frac{e^{-k\frac{\|u-v\|^2}{2t}}}{(2t\pi/k)^{k/2}}(1 + O(\frac{k\|u-v\|}{\sqrt{t}}))$$

5 The trace formula for lattice graphs

Let L denote a subset of $\mathbb{Z}^k$ and we write

$$\langle L\rangle = \{\sum_x n_x x : x \in L, n_x \in \mathbb{Z}\}$$

Let P^k/L denote the graph formed from P^k by identifying vertices x and $x+y$ in $\langle L\rangle$ for any y. Furthermore, we define

$$L^* = \{y \in (\mathbb{R}/\mathbb{Z})^k : \langle x, y\rangle \in \mathbb{Z} \text{ for all } x \in L\}$$

We will prove the following:

Theorem 14 *(Discrete Poisson Formula)*

$$|P^k/L| \sum_{y \in <L>} K_t(y) = \sum_{x \in L^*} exp(-\frac{2t}{k} \sum_{j=1}^{k} \sin^2 \pi x_j))$$

where $K_t(y)$ is equal to the heat kernel $\mathbf{H}_t(x, x+y)$ of an infinite path and $x = (x_1, \ldots, x_k)$.

Proof: To prove the trace formula for the graph P^k/L, the heat kernel $H(x,x)$ are all equal and

$$H(x,x) = \sum_{y \in \langle L \rangle} K_t(y)$$

where K is defined by the heat kernel of P^k.

We need to show that the eigenvalues of P^k/L is exactly $\frac{2}{k} \sum_{j=1}^{k} \sin^2 \pi x_j$ for $x = (x_1, \ldots, x_k)$ in L^*. This can be shown by considering the eigenfunction f_x for $x = (x_1, \ldots, x_k)$ in L^*, defined as follows:

$$f_x(y) = e^{2\pi i <x,y>} \text{ for } y = (y_1, \ldots, y_k) \in P^k, y_i \in \mathbb{Z}$$

It is straightforward to check that

$$\mathcal{L} f_x(y) = \frac{1}{2k} \sum_{j=1}^{k} (2 - 2\cos 2\pi x_j) f_x(y)$$

Example 1: For $k = 1$ and $L = \{n\}$, P/L is just an $n-$cycle and we get

$$\sum_{k=0}^{\infty} \sum_{j=-\infty}^{\infty} (-1)^{nj} \frac{\binom{2k}{k+nj}(-\frac{t}{2})^k}{k!} = \frac{1}{n} \sum_{k=0}^{n-1} e^{-2t \sin^2 \frac{\pi k}{n}}$$

Example 2: For $k = 2$ and $L = \{(3,5),(4,0)\}$, we have

$$\sum_{k=0}^{\infty} \sum_{k'=0}^{\infty} \sum_{j=-\infty}^{\infty} \sum_{j'=-\infty}^{\infty} \binom{2k}{k+3j+4j'} \binom{2k'}{k'+5j} \frac{(-\frac{t}{4})^{k+k'}}{k!\ k'!} = \frac{1}{20} \sum_{k=0}^{3} \sum_{k'=0}^{4} e^{-t(\sin^2 \frac{k\pi}{4} + \sin^2 \frac{(k' - \frac{3k}{4})\pi}{5})}$$

Example 3: For $k = 3$ and $L = \{(1,1,1)\}$, P^3/L is the triangular lattice and we have

$$\sum_{a+b+c=0} \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \sum_{k_3=0}^{\infty} (-1)^{a+b+c} \binom{2k_1}{k_1+a} \binom{2k_2}{k_2+b} \binom{2k_3}{k_3+c} \frac{(-\frac{t}{6})^{k_1+k_2+k_3}}{k_1!k_2!k_3!}$$
$$= \frac{1}{\pi} \int_0^{\pi} \int_0^{\pi} e^{-\frac{2t}{3}(\sin^2 x + \sin^2 y + \sin^2(x+y))} dxdy$$

References

[1] B. Bollobás, Extremal Graph Theory, Academic Press, London (1978).

[2] R. Brooks, The spectral geometry of k-regular graphs, *Journal D'analyse Mathématique*, 57 (1991) 120-151.

[3] F. R. K. Chung and S. -T. Yau, Eigenvalues of graphs and Sobolev inequalities, *Combinatorics, Probability and Computing*, 4 (1995) 11-26.

[4] F. R. K. Chung and S. -T. Yau, A Harnack inequality for homogeneous graphs and subgraphs, *Communications in Analysis and Geometry*, 2, (1994), 628-639.

[5] F. R. K. Chung and S. -T. Yau, Coverings of graphs and their heat kernels, preprint.

[6] F. R. K. Chung, *Spectral Graph Theory*, CBMS Lecture Notes, 1996 , AMS Publication.

[7] S. T. Yau and Richard M. Schoen, *Lectures on Differential Geometry*, (1994), International Press, Cambridge, Massachusetes

[8] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, 1995.