## Section 1: Introduction to the Heat Equation

# 1 Definitions

**PDE.** Partial differential equation, an equation or system of equations that
determines a function or collection of functions, called the *solution*. The
solution is a function of more than one variable. The equations involve
partial derivatives of the unknown solution.

**Data.** Parameters, and especially functions, needed to determine the solution.

**Existence.** There is a solution to the PDE.

**Uniqueness.** There is only one solution corresponding to a given set of data.
Often with side conditions on the type of the solution – a unique *bounded*
solution, or *positive* solution, or *viscosity* solution (details of this in a later
class).

**Well posedness.** Existence + uniqueness (the PDE together with the data
determine exactly one solution) + the solution depends continuously on
the data. Changing the data less than $\delta$ changes the solution less than $\epsilon$,
in some sense.

**Strong solution.** A function such that all the partial derivatives mentioned in
the PDE exist in the ordinary sense of calculus and these partial deriva-
tives make the PDE true at every point at which it applies.

**Classical solution.** The same as strong solution.

**Weak solution.** A function that may not be differentiable but satisfies con-
ditions that were used to derive the PDE. Solutions of Hamilton Jacobi
equations in control theory are common examples.

**General theory.** There isn't one. Nothing that applies to all PDEs seems to
be very interesting.

**Model problem.** A simple PDE whose solutions can be understood in detail,
often by direct calculation. Of interest because it suggests what solutions
of related but more difficult PDEs should look like. Often used as counter-
examples to questions like existence and well posedness.

**Space and time variables.** The solution is a function of a variable, $t$, that represents time, and one or more other *spatial* variables, $x$, that represent other things. Spatial variables need represent coordinates or positions in physical space. For example, in finance, the space variables may represent the sized of investments in various instruments.

**Evolution equation.** A PDE that models the change of a solution through the passage of time.

**Steady state.** A PDE or solution to a PDE that does not depend on a time variable. Sometimes found as the limit as $t \to \infty$ or $t \to -\infty$ of the solution of an evolution equation.

**Initial or final data.** Data that specify the values of the solution or derivatives of the solution at particular values of $t$.

**Boundary contidions.** Data that specify the values of the solution or derivatives of the solution at particular values of $x$.

## 2    The heat equation

The solution is a function, $u(x,t)$, that depends on one or more space variables and one time variable. It represents the temperature (say, in a block of material or still air) at location $x$ at time $t$. If there are $n$ space variables, we write $x = (x_1, \ldots x_n)$. The $x_k$ are the *coordinates* or the *components* of $x$. The *Laplace operator*, or the *Laplacian*, of a function of $n$ variables is

$$\triangle f \;=\; \partial_{x_1}^2 f + \cdots + \partial_{x_n}^2 f \;.$$

For example, if $f(x) = |x|^2 = x_1^2 + \cdots + x_n^2$, then $\triangle f = 2 + \cdots + 2 = 2n$. *The heat equation* is

$$\partial_t u \;=\; \frac{1}{2}\, \triangle u \;. \tag{1}$$

For example, $u(x,t) = |x|^2 + nt$ satisfies the heat equation. The heat equation is an evolution equation that determines the temperature distribution at later times (larger $t$) from the temperature distribution at an earlier time, and possibly also from boundary conditions.

Here is one textbook derivation of the heat equation. $u(x,t)$ is the energy density (actually, the difference between the energy density at a given point and a background energy density. $u = 0$ does not mean zero energy, but zero energy difference) at location $x$ at time $t$, measured in energy per unit volume (say, ergs per cubic centimeter). The specific energy density is energy per unit mass, $e$, say in ergs per gram. This is the energy density divided by the mass density, $\rho$:

$$u \;=\; \frac{\text{energy}}{\text{volume}} \;=\; \frac{\text{energy}}{\text{mass}} \times \frac{\text{mass}}{\text{volume}} \;=\; e \cdot \rho \;.$$

2

The specific heat, $\kappa$, is the relation between the specific energy density and the temperature. It is the amount of energy (per unit mass) needed to raise the temperature by one degree: $e = \kappa T$ (so that $\Delta T = 1$ corresponds to $\Delta e = \kappa$). We assume that temperature and energy differences are small enough that the relationship between temperature (temperature difference, actually) and specific energy is linear. We normalize so that $T = 0$ is the background temperature corresponding to the background energy difference $e = 0$.

Suppose $V$ is a volume (region) in space with boundary surface $\Gamma$. For example, in three dimensions, $V$ could be a *ball* consisting of points within a distance $R$ of a center: $|x - x_c| \leq R$. Then $\Gamma$ would be the corresponding *sphere* of points with $|x - x_c| = R$. In one dimension, this $V$ would be an interval $[x_c - R, x_c + R]$ and the corresponding $\Gamma$ would be the endpoints $\{x_c - R, x_c + R\}$. The total energy within $V$ at a given time, also called the heat, is

$$H_V(t) = \int_V u(x,t)\, dx \,,$$

where the integral is over a one or two or three dimensional volume, depending. The *heat flux* is a vector field, $F(x,t)$, describing the flow of heat energy. If $\Gamma$ is a surface and $x \in \Gamma$ is a point on the surface, then we write $n(x)$ for a unit normal vector normal to $\Gamma$ at $x$. $\Gamma$ should be *oriented*, which means that has two distinct sides or faces (the Möbius band is a surface that has only one side), and the normals $n(x)$ all point toward the same side. The heat flux across $\Gamma$ is

$$\int_\Gamma F(x) \cdot n(x)\, dA(x) \,,$$

where $dA(x)$ refers to integration with respect to the surface area of $\Gamma$ (arc length in two dimensions and a sum over boundary points in one dimension.

These are all just definitions. The two physical hypothesis lead to the heat equation. The first is local conservation of heat energy: the only way $H_V(t)$ can change is by heat flowing across the boundary. This is not true, for example, if heat is produced within $V$ by chemical reactions. The mathematical statement is

$$\partial_t H_V(t) + \int_\Gamma F(x,t) \cdot n(x)\, dA(x) = 0 \,,$$

where the normals are oriented away from the interior of $V$. The second hypothesis is the *Fourier law*, which states that $F(x,t)$ is proportional to the negative of the temperature gradient: $F(x,t) = -\mu \nabla T(x,t)$. This states that heat energy flows from high temperature towards low temperature at a rate proportional to the temperature gradient.

The derivation returns to mathematics. Making substitutions, we get $F(x,t) = -D\nabla u(x,t)$, where the *diffusion coefficient* is $D = \mu/\kappa\rho$. The conservation equation then becomes

$$\partial_t \int_V u(x,t)\, dx = D \int_\Gamma \nabla u(x,t) \cdot n(x)\, dA(x) \,.$$

On the left, we can move the time derivative inside the integral. On the right, we can apply the divergence theorem, remembering that

$$\mathrm{div}\bigl(\nabla u\bigr) \;=\; \partial_{x_1}\partial_{x_1}u + \cdots + \partial_{x_n}\partial_{x_n}u \;=\; \triangle u \;.$$

The result is that for any $V$,

$$\int_V \partial_t u(x,t)\, dx \;=\; D \int_V \triangle u(x,t)\, dx \;.$$

If $\partial_t u$ and $\triangle u$ are continuous, this implies that $\partial_t u(x,t) = D \triangle u(x,t)$ for all relevant values of $x$ and $t$. (If, say $\partial_t u(x,t) > D \triangle u(x,t)$ at some $x$ and $t$, then the same is true in a sufficiently small ball, $V$, about $x$, which makes the integrals also unequal.) The final step is to *rescale* time, which means choosing a new time variable $t' = 2Dt$, so that $2D\partial_{t'}u = \partial_t u$ and $\partial_{t'}u = \frac{1}{2}\triangle u$. This is *the* heat equation (1), if we ignore the distinction between $t'$ and $t$. We chose to rescale time so as to leave a factor $\frac{1}{2}$ on the right of the heat equation. Another common choice is $t' = D^{-1}t$, which does not have the $\frac{1}{2}$. Both normalizations are convenient in some ways and inconvenient in others.

For mathematical analysis we assume at first that the heat equation (1) applies in all of space. The heat equation is an evolution equation. The heat distribution at time $t$ is the function of $x$ given by $u(x,t)$. We will see that the heat equation allows us to determine the heat distribution at time $t_2$ from the distribution at time $t_1$ only if $t_2 \geq t_1$ (the case $t_2 = t_1$ being particularly easy). More tecnically, evolving the heat equation forward in time is a well posed problem, while evolving it backward in time is not well posed (as shown in Section 2).

There are many physical situations where the heat equation is an accurate quantitative description. There are others where it is a poor quantitative description but a reasonable qualitative one. The whole class of *linear parabolic* PDEs is well posed for evolving in one direction but not the other and shares smoothing and other qualitative properties with the heat equation. This makes the heat equation important as a model problem for understanding what is and is not possible with diffusion equations in general.

The heat equation also is a model of *diffusion*. For example, suppose there is a lump of sugar in the bottom of a glass of warm water. The sugar diffuses through the water. We can picture individual molecules of sugar wandering randomly through the water. If $u(x,t)$ is the concentration of sugar at time $t$ at location $x$ in the glass, then $u$ may satisfy the heat equation once time is properly rescaled (if the water is still and the sugar concentration is not too large and ...). In this setting, the Fourier law is called *Fick's law*. It says that the net flux (vector) of sugar is proportional to the negative of gradient of the sugar density. Again, without chemical reactions, the amount of sugar in a *control volume*, $V$, changes only by sugar crossing the boundary. This is a *net* flux. Sugar molecules wander around randomly. Those that start in low density regions may wander to regions of higher density. But these are outnumbered by particles starting in high density regions wandering to low density regions,

because particles starting in high density regions outnumber those starting in low density regions.

The same reasoning applies to probability density. If we think of the sugar molecules wandering randomly through the water, the probability density for an individual molecule should be proportional the the actual number density of sugar molecules. Again there is a probability flux vector that points in the direction of lower probability.

# 3  Special solutions

One of the things that makes the heat equation useful as a model problem is that it has many explicit solutions. For example, there are polynomial solutions. The simplest polynomials are constants and linears, such as $u(x,t) = v \cdot x$ for any vector $v$ (note that $\triangle u = 0$). We noted the quadratic solution $|x|^2 + nt$. There are more in the homework.

Other solutions have the product form $u(x,t) = f(x)g(t)$, which is called *separation of variables*. A substitution shows that

$$\frac{1}{g(t)}\partial_t g(t) \;=\; \frac{1}{f(x)}\frac{1}{2}\triangle f(x) \;=\; \lambda \;.$$

We argue that $\lambda$ is a constant independent of $x$ and $t$. For example, it is independent of $x$ because $(\partial_t g(t))/g(t)$ is independent of $x$. Therefore, we have $\partial_t g(t) = \lambda g(t)$, so $g(t) = Ce^{\lambda t}$. Similarly, $\frac{1}{2}\triangle f = \lambda f$. There are many solutions of this, but one important one is $f(x) = e^{v \cdot x}$, where $\frac{1}{2}|v|^2 = \lambda$. For example, in one dimension, there is a solution of the form

$$u(x,t) \;=\; e^{3x+(9/2)t} \;. \tag{2}$$

A particularly useful separation of variables solution has a purely imaginary spatial exponent: $f(x) = e^{ip \cdot x}$, where $p$ is real. The corresponding solution is

$$u(x,t) = Ce^{ip \cdot x}e^{-|p|^2 t/2} \;. \tag{3}$$

We can understand the complex exponential solutions using the ancient formula $f(x) = e^{ip \cdot x} = \cos(p \cdot x) + i\sin(p \cdot x)$. This shows that $f$ varies only if we move in the $p$ direction and is constant if we move in a direction perpendicular to $p$. Moreover, $f$ is periodic. If we move a distance $2\pi/|p|$ in the $p$ direction, $f$ repeats itself. We call $p$ the *wave vector* (particularly in one dimension, $p$ or $|p|$ may be called the wave *number*.) and $2\pi/|p|$ the *wave length*. $p$ (or $|p|$ also is called the *frequency*, or *spatial frequency*. Large wave vector or high frequency corresponds to short wave length, and small wave number corresponds to long waves.

The diffusion picture gives a qualitative interpretation of these special solutions. For this, we consider $u(x,t)$ to be the difference between the actual density and a constant background density. $u > 0$ means that the local density

is higher than the background density. while $u < 0$ means the local density is lower. The solution $u = |x|^2 + nt$ is a constantly rising tide. Sugar is rushing from regions where $|x|$ is large (large $|x|$) to regions where $|x|$ is smaller. This results in the value in the center, $u(0,t)$ growing linearly in time. More quantitatively, the sugar flux is $F(x,t) = -\frac{1}{2}\nabla |x|^2 = -x$, which is a vector pointed toward the origin whose magnitude is proportional to the distance from the origin (Duh!). If $V$ is a ball of radius $R$ in three dimensions centered at $x = 0$, then the flux vector is normal to $\Gamma$ and the total flux into $V$ is surface area $(4\pi R^2)$ times the magnitude of the flux vector $(R)$, so the total flux is $4\pi R^3$. Therefore, the change in total sugar within $V$ between time $t = 0$ and time $t_1$ is $4\pi R^3 t$. We calculate the same change as

$$
\int_{|x|\leq R} u(x,t_1)\, dx \;-\; \int_{|x|\leq R} u(x,0)\, dx
$$
$$
= \;\int_{|x|\leq R} \left( |x|^2 + 3t \right) dx \;-\; \int_{|x|\leq R} |x|^2\, dx
$$
$$
= \;\frac{4}{3}\pi R^3 \cdot 3t \;=\; 4\pi R^3 t \;,
$$

which agrees with the flux calculation.

The real exponential solution is consistent with the idea that the density is increasing because of sugar diffusing from the higher density regions to the right (if $\lambda > 0$). Consider the real part of the complex exponential solution (3), $u(x,t) = \cos(p \cdot x)e^{-|p|^2 t/2}$. Sugar diffuses from positive to negative density regions, causing the maximum deviation $(\max - \min = 2e^{-|p|^2 t/2})$ to decrease exponentially. The (exponential) decay rate is faster for larger $p$ because the separation between positive and negative, which is the distance they have to diffuse to cancel, is (half) the wavelength, $\pi/|p|$.

The complex exponential solution illustrates the relationship between time and space scales in the heat equation and other diffusion problems. The length scale is the wavelength: $L \sim 1/|p|$. We don't worry about the possible $2\pi$ factor because we are talking about qualitative orders of magnitude rather than precise distances. The time scale, $T$, is determined by the exponential decay rate: $e^{-t/T}$. Comparing to (3) gives $T \sim 1/|p|^2 \sim L^2$. Thus, the time scale is the square of the length scale. If we had not nondimensionalized $(t' = 2Dt)$, the relation would have been $DT \sim L^2$. A larger diffusion coefficient makes diffusion faster – the same $L$ and a larger $D$ makes a smaller $T$.

Another extremely useful explicit solution is the *fundamental solution*, or *Green's function*:

$$
u(x,t) \;=\; \frac{1}{(2\pi t)^{n/2}} e^{-|x|^2/2t} \;. \tag{4}
$$

For example, in one dimension both $\partial_t \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}$ and $\frac{1}{2}\partial_x^2 \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}$ are equal to

$$
\left( \frac{-1}{2t} + \frac{x^2}{2t^2} \right) \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} \;.
$$

6

This solution corresponds to a high density blob diffusing away from the origin. The maximum density is $u(0,t) = 1/\left(2\pi t\right)^{n/2}$, which decreases from an infinite value when $t = 0$ towards zero as $t \to \infty$. If $x \neq 0$, then $u(x,t) \to 0$ as $t \to 0$, i.e. there is no sugar anywhere but $x = 0$ when $t = 0$. The total mass for this fundamental solution, the total amount of sugar, is

$$\int_{R^n} u(x,t)\, dx \;=\; 1\;,$$

for any $t > 0$, and for $t = 0$ in particular. Therefore we write $u(x,0) = \delta(x)$ and call it a *point mass* located at the origin.

Another feature of the fundamental solution is that it has *radial symmetry*. This is because every direction in space is the same as every other direction. In Fick's law, the diffusion flux is proportional to the density gradient with a proportionality constant that does not depend on the direction of the gradient. Therefore, the solution corresponding to sugar diffusing away from a single point a the origin should depend only on the distance from the origin (and time), but not on the direction from the origin.

We could start at time $t = 0$ with a point mass at a point $y \neq 0$. The corresponding solution would be

$$G_0(y,x,t) \;=\; u(x-y,t) \;=\; \frac{1}{\left(2\pi t\right)^{n/2}} e^{-|x-y|^2/2t}\;. \tag{5}$$

This is the *free space Green's function* for the heat equation. The subscript 0 implies that is a starting point from which Green's functions for other problems may be expressed.

## 4   Superposition

The heat equation has the *superposition principle*. If $u_1(x,t)$ and $u_2(x,t)$ are solutions, then $u(x,t) = u_1 + u_2$ also is a solution. PDEs with this property are called *linear*. It is important to be careful with the definition of linearity. The American style option pricing model of Black and Scholes uses a PDE that looks linear, but the whole problem does not satisfy the superposition principle because of nonlinear free boundary conditions. The real test of linearity is the superposition principle. By the way, in quantum mechanics superposition is a fundamental principle of nature. In general PDE it is not really a principle, but a property that sometimes holds and sometimes does not. Nobody calls it the superposition property, however. Similarly, the so called maximum principle holds for some PDEs but not for others.

Of course we can add more than one solution and use coefficients other than one. If $u_k(x,t)$ are solutions and the numbers $w_k$ are arbitrary constants, then $u(x,t) = \sum_k w_k u_x(x,t)$ also is a solution. For example, suppose we start with point masses at locations $y_k$ with masses $w_k$, then the total density would be

the sum of the densities corresponding to the individual masses:

$$u(x,t) \; = \; \sum_k w_k \, G_0(x, y_k, t) \; .$$

We also can have integral superpositions. We describe the situation with an abstractly at first. Suppose that $\theta$ is a continuous parameter and for each $\theta$, $u(x, t, \theta)$ is a solution of the heat equation. Suppose the $w(\theta)$ is some continuous *weight function*. It is easily checked that the integral linear superposition,

$$u(x,t) \; = \; \int w(\theta) u(x, t, \theta) \, d\theta \; , \tag{6}$$

also is a solution, if the integral converges well enough.

The *free space initial value problem* is to find the solution of the heat equation that also satisfies initial conditions $u(x, 0) = f(x)$. We seek families of special solutions $u(x, t, \theta)$ that are rich enough so that any solution of the initial value problem may be represented in the form (6). This may be done using the Green's functions (5). Here $y$ corresponds to the abstract parameter, $\theta$. The simple formula that corresponds to (6) has $w(y) = f(y)$, so

$$u(x,t) \; = \; \int_{R^n} f(y) G_0(y, x, t) \, dy \; . \tag{7}$$

We give an intuitive derivation of (7) in one dimension. It is easy but more cumbersome to give the same derivation in more dimensions. The idea is to choose a small spacing, $h$, and put small point masses at the *lattice points* (or *grid points*, or *mesh points*) $y_k = kh$. Suppose that the bin $B_k$ is the interval of $x$ values closest to $y_k$: $B_k = [y_k - \frac{h}{2}, y_k + \frac{h}{2}]$. The mass in the initial density distribution, $f$ in $B_k$ is

$$\int_{B_k} f(x) \, dx \; \approx \; h f(y_k) \; = \; w_k \; ,$$

at least if $f$ is continuous. Let us approximate the initial distribution by a collection of point masses with the appropriate the appropriate masses:

$$f(x) \; \approx \; \sum_k w_k \delta(x - x_k) \; , \tag{8}$$

This may seem to be a (literally) rough approximation, but with the point masses close together, it should not take long for diffusion to make diffusing particles to forget exactly which mass they came from. That is, we expect the density even from a sum of point masses to become reasonably smooth after just a short time. This turns to happen. Using the superposition principle, the solution corresponding to the right side of (8) is

$$u(x,t,h) \; = \; \sum_k w_k G_0(x - y_k) \; = \; \sum_k h f(y_k) G_0(x - y_k) \; .$$

This sum converges to the integral (7) as $h \to 0$.

It is not hard to give mathematical proofs of the following, under the hypothesis that $f$ is continuous and bounded.

- The integral (7) converges for any $x$ if $t > 0$.

- For any $t > 0$ all mixed partial derivatives of $u$ of any order exist, and may be computed by differentiating under the integral sign.

- In particular, $u(x,t)$ satisfies the heat equation (in the strong classical sense), because $G_0(y,x,t)$ satisfies the heat equation for each value of the parameter $y$.

- For each $x$, $u(x,t) \to f(x)$ as $t \downarrow 0$ ($t \downarrow 0$ means that $t$ approaches zero through positive values, or goes to zero *from above*.).

To understand the last point, note that for each $x$, the function $G_0(y,x,t)$ is a probability density (as a function of $y$, look to (5) to confirm this) that is strongly localized about $x$. Therefore, the integral (7) gives $u(x,t)$ as an average of of values of $f(y)$ for $y$ near $x$. If $f$ is continuous, all these $f(y)$ values should be close to $f(x)$, which implies that $u(x,t)$ should be close to $f(x)$. The mathematical subtlety (not a huge one) is that $G_0$ is non-zero (though extremely small) even for values of $y$ far from $x$. Therefore distant $f(y)$ values make some contribution to $u(x,t)$. If $f$ is bounded, these contributions vanish as $t \downarrow 0$.

The bullet points above, taken together, imply that (7) provides a complete general solution of the initial value problem for the heat equation in free space. In turn, the representation formula (7) gives much information about the solution. One thing it does not do is prove uniqueness. Is there another function, $v(x,t) \neq u(x,t)$, that also satisfies the heat equation and the initial conditions? This question in Section 3 and in Section 2 of Bob Kohn's notes.

## 5  Cancellation

The heat equation, as with other diffusion equations, has the feature that positive and negative "heat" can diffuse from different places and cancel each other. Suppose $u_1(x,t)$ and $u_2(x,t)$ satisfy the heat equation with initial data $u_1(x,0) = f_1(x)$ and $u_2(x,0) = f_2(x)$. Suppose further that $f_1$ is the *positive part* of the initial data and $f_2$ the negative part: $f_1(x) \geq 0$ and $f_2(x) \leq 0$ for all $x$. The Green's function representations of $u_1$ and $u_2$ given by (7) imply that $u_1(x,t) > 0$ and $u_2(x,t) < 0$ for all $x$ and $t > 0$. Let $u(x,t) = u_1(x,t) + u_2(x,t)$ be the solution of the heat equation with initial data $f(x) = f_1(x) + f_2(x)$. Cancellation is the fact that

$$|u(x,t)| = |u_1(x,t) + u_2(x,t)| < |u_1(x,t)| + |u_2(x,t)| .$$

Cancellation is the cause of the exponential decay of the complex exponential solutions. The mass in the positive and negative parts of $\cos(p \cdot x)$ have the same mass so they can cancel more and more completely as time goes on.

We also can calculate the cancellation of a positive and negative point mass as they diffuse together. We do this in one dimension for simplicity. The initial data $f(x) = \delta(x-1) - \delta(x+1)$ corresponds to a positive unit mass at $x = 1$ and a negative unit mass at $x = -1$. The superposition principle gives the solution as

$$u(x,t) = \frac{1}{\sqrt{2\pi t}} \left( e^{-(x-1)^2/2t} - e^{-(x+1)^2/2t} \right).$$

The total "uncancelled" mass at time $t = 0$ is $\left| \int f(x) \right| dx = 2$. The total mass remaining uncancelled mass that remains at time $t$ is $M(t) = \int |u(x,t)| \, dx$. By symmetry (see Kohn's notes, Section 2, page 5), this is given by

$$M(t) = 2 \int_0^\infty u(x,t) \, dx = \frac{1}{\sqrt{2\pi t}} \int_0^\infty \left( e^{-(x-1)^2/2t} - e^{-(x+1)^2/2t} \right) dx.$$

We approximate the integral on the right for large $t$ by expanding the exponentials as:

$$e^{-(x-1)^2/2t} = e^{-x^2/2t} \, e^{x/t} \, e^{-1/2t} \approx e^{-x^2/2t} \left( 1 + \frac{x}{t} \right). \tag{9}$$

With this,

$$\int_0^\infty \left( e^{-(x-1)^2/2t} - e^{-(x+1)^2/2t} \right) dx \approx \int_0^\infty e^{-x^2/2t} \frac{2x}{t} \, dx = 2.$$

Therefore, $M(t) \approx 4/\sqrt{2\pi t} \sim t^{-1/2}$ for large $t$.

A detailed rigorous justification of the approximation (9) is too much of a distraction to present here. Instead, use the informal idea that important $x$ values have the order of magnitude $\sqrt{t}$. Then the kept term has the order of magnitude $x/t \sim 1/\sqrt{t}$, while the neglected terms have magnitude $x^2/t^2 \sim 1/t$ and $1/t$ itself.