Chapter 6

Wave equations on $\mathbb{R}^n$

Solutions of this equation describe the propagation of light, of sound waves in a gas or a fluid, of gravitational waves in the interstellar vacuum, and many other phenomena. It is one of my favourite equations. Posed in $\mathbb{R}^1_t \times \mathbb{R}^n_x$ the initial value problem, (or *Cauchy problem*), for the equation looks very similar to (2.10) of Chapter 2;

$$\partial_t^2 u - \Delta u = 0, \qquad x \in \mathbb{R}^n, t \in \mathbb{R}^1 \ , \tag{6.1}$$

with initial or Cauchy data for $u(x,t)$ given by

$$u(0,x) = f(x) \ , \qquad \partial_t u(0,x) = g(x) \ .$$

The given initial data $u(0,x) = f(x)$ is often referred to as the initial position or displacement of the field $u(t,x)$, while the data for $\partial_t u(0,x) = g(x)$ is called the initial momentum or velocity.

6.1. Wave propagator by Fourier synthesis

Assuming first that $f(x), g(x) \in \mathcal{S}$, we may take the Fourier transform of the equation to obtain

$$\partial_t^2 \hat{u}(\xi,t) + |\xi|^2 \hat{u}(\xi,t) = 0 \ . \tag{6.2}$$

Solutions of this second order ODE are composed of linear combinations of $e^{\pm i|\xi|t}$; taking into account that $\hat{u}(\xi,0) = \hat{f}(\xi)$ and $\partial_t \hat{u}(\xi,0) = \hat{g}(\xi)$, we derive the expression

$$\hat{u}(t,\xi) = \cos(|\xi|t)\hat{f}(\xi) + \frac{\sin(|\xi|t)}{|\xi|}\hat{g}(\xi) \ . \tag{6.3}$$

The inverse Fourier transform gives the solution

$$u(t,x) = \frac{1}{\sqrt{2\pi}^n} \int e^{i\xi\cdot x} \left(\cos(|\xi|t)\hat{f}(\xi) + \frac{\sin(|\xi|t)}{|\xi|}\hat{g}(\xi) \right) d\xi \ . \tag{6.4}$$

we observe from (6.3) that $\hat{u}(t,\xi) \in \mathcal{S}(\mathbb{R}^n_x)$ for each time t, and therefore our solution given in (6.4) is Schwartz class as well. However this high level of smoothness is not at all necessary, and one notes that the expression (6.4) makes sense whenever $\hat{f}(\xi), \frac{\hat{g}(\xi)}{|\xi|} \in L^2(\mathbb{R}^n_x)$.

Theorem 6.1. *For $f, g \in \mathcal{S}$, the expression* (6.4) *gives a solution of the wave equation $u(t,x) \in \mathcal{S}(\mathbb{R}^n_x)$ for each time $t \in \mathbb{R}$. For $\hat{f}(\xi), \frac{\hat{g}(\xi)}{|\xi|} \in L^2(\mathbb{R}^n_x)$, then* (6.4) *gives a weak solution to the wave equation, in the sense that* (6.2) *is satisfied.*

One basic property satisfied by solutions of the wave equation is the principle of 'energy' conservation under time evolution. The energy of a solution is given by

$$E(u) = \frac{1}{2} \int (\partial_t u(t,x))^2 + |\nabla_x u(t,x)|^2 \, dx \ . \tag{6.5}$$

Theorem 6.2. *For $(f,g) \in H^1(\mathbb{R}^n_x) \times L^2(\mathbb{R}^n_x)$ the energy* (6.5) *is conserved for solutions of the wave equation* (6.1).

Proof. Using the expression (6.3) for the Fourier transform of the solution, we give a similar formula for $(\partial_x u(t,x), \partial_t u(t,x))^T$

$$\begin{pmatrix} \partial_x u(t,x) \\ \partial_t u(t,x) \end{pmatrix} = \frac{1}{\sqrt{2\pi}^n} \int e^{i\xi\cdot x} \begin{pmatrix} \cos(|\xi|t) & \frac{i\xi}{|\xi|}\sin(|\xi|t) \\ \frac{i\xi}{|\xi|}\sin(|\xi|t) & \cos(|\xi|t) \end{pmatrix} \begin{pmatrix} i\xi\hat{f}(\xi) \\ \hat{g}(\xi) \end{pmatrix} d\xi \ .$$

By inspection this is a well defined vector values function in $[L^2(\mathbb{R}^n)]^{n+1}$ as long as $(\nabla_x f(x), g(x))^T \in [L^2(\mathbb{R}^n)]^{n+1}$, which it is by hypothesis. By the Plancharel identity

$$\begin{aligned} \|\partial_t u(x,t)\|^2_{L^2} + \|\partial_x u(x,t)\|^2_{L^2} &= \| \begin{pmatrix} \cos(|\xi|t) & \frac{i\xi}{|\xi|} \\ \frac{i\xi}{|\xi|}\sin(|\xi|t) & \cos(|\xi|t) \end{pmatrix} \begin{pmatrix} i\xi\hat{f}(\xi) \\ \hat{g}(\xi) \end{pmatrix} \|^2_{L^2} \\ &= \| \begin{pmatrix} i\xi\hat{f}(\xi) \\ \hat{g}(\xi) \end{pmatrix} \|^2_{L^2} = \|\nabla_x f\|^2_{L^2} + \|g\|^2_{L^2} \ . \end{aligned}$$

The second to last equality holds because the 2×2 matrix featured in this calculation is unitary. □

An alternative proof works in the case that we also have $\partial_t^2 u, \Delta u \in L^2$, then for solutions of (6.1);

$$\frac{d}{dt}E(u) = \tfrac{1}{2}\int 2\partial_t u \partial_t^2 u + 2\partial_x u \partial_t \partial_x u \, dx$$
$$= \int \partial_t u \left(\partial_t^2 u - \Delta u\right) \, dx = 0 \ .$$

6.2. Lorentz transformations

Just as the Laplace operator $\Delta = \sum_{j=1}^n \partial_{x_j}^2$ is invarialnt under translations $x' = x + c$ and rotations $x' = Rx$, where $R^T = R^{-1}$, the wave operator or d'Alembertian

$$\Box = \partial_t^2 - \Delta = \partial_t^2 - \sum_{j=1}^n \partial_{x_j}^2$$

is invariant under a group of transformations of the space-time $\mathbb{R}_t^1 \times \mathbb{R}_x^n$ known as the Lorentz group. Elements of this transformation group are generated by the same rotations of $\mathbb{R}_x^n$ as above, and *hyperbolic rotations* which involve time as well as space. Define a hyperbolic rotation in the (t, x_1) coordinate plane in $\mathbb{R}^2 \subseteq \mathbb{R}^{n+1}$ by

$$\begin{pmatrix} t' \\ x'_1 \end{pmatrix} = \begin{pmatrix} \cosh(\psi) & \sinh(\psi) \\ \sinh(\psi) & \cosh(\psi) \end{pmatrix} \begin{pmatrix} t \\ x_1 \end{pmatrix} = H(\psi) \begin{pmatrix} t \\ x_1 \end{pmatrix} \ .$$

One calculates that $\det(H(\psi)) = \cosh(\psi)^2 - \sinh(\psi)^2 = 1$ and that $H^{-1}(\psi) = H(-\psi)$. Vector fields in the (t, x_1) coordinate plane trasnsform as follows

$$\begin{pmatrix} \partial_t \\ \partial_{x_1} \end{pmatrix} = \begin{pmatrix} \cosh(\psi) & \sinh(\psi) \\ \sinh(\psi) & \cosh(\psi) \end{pmatrix} \begin{pmatrix} \partial_{t'} \\ \partial_{x'_1} \end{pmatrix} = H(\psi) \begin{pmatrix} \partial_{t'} \\ \partial_{x'_1} \end{pmatrix} \ . \tag{6.6}$$

The Lorentz group of transformations on the space-time $\mathbb{R}_t^1 \times \mathbb{R}_x^n$ is generated by all spatial rotations R and by the hyperbolic rotations $H(\psi)$. It is a Lie group known as $SO(1, n)$. The set of space-time translations $(t', x') = (t + b, x + c)$ leave the wave equation invariant, and so do the elements of the Lorentz group.

Proposition 6.3. The Lorentz transformations leave the d'Alembertian operator invariant.

Proof. Since the group generators involving translations obviously leave the d'Alembertian invariant, and all spatial rotations R just involve the Laplacian, it suffices to check invariance under the hyperbolic rotations $H(\psi)$. Write the d'Alembertian operator using matrix notation

$$\Box u = (\partial_t^2 - \Delta) u = \begin{pmatrix} \partial_t \\ \partial_{x_1} \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & -I_{n\times n} \end{pmatrix} \begin{pmatrix} \partial_t \\ \partial_{x_1} \end{pmatrix} u \ .$$

Use the transformation rule (6.6) for vector fields to check the invariance of the d'Alembertian

$$\begin{pmatrix} \partial_{t'} \\ \partial_{x'_1} \end{pmatrix}^T \begin{pmatrix} H_{2\times 2}^T & 0 \\ 0 & I_{(n-1)\times(n-1)} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -I_{n\times n} \end{pmatrix} \begin{pmatrix} H_{2\times 2} & 0 \\ 0 & I_{(n-1)\times(n-1)} \end{pmatrix} \begin{pmatrix} \partial_{t'} \\ \partial_{x'_1} \end{pmatrix} .$$

The only nontrivial part of this matrix calculation is the upper right hand 2×2 block

$$\begin{aligned} H_{2\times 2}^T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} H_{2\times 2} &= \begin{pmatrix} \cosh(\psi) & \sinh(\psi) \\ \sinh(\psi) & \cosh(\psi) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cosh(\psi) & \sinh(\psi) \\ \sinh(\psi) & \cosh(\psi) \end{pmatrix} \\ &= \begin{pmatrix} \cosh^2(\psi) - \sinh^2(\psi) & 0 \\ 0 & \sinh^2(\psi) - \cosh^2(\psi) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \end{aligned}$$

which is precisely to say that the d'Alembertian $\Box$ is invariant under $H(\psi)$. This is the stated result. $\Box$

The effects of the Lorentz transformation $H(\psi)$ on the time axis $\{x = 0\}$ are given by

$$x'_1 = \sinh(\psi)t \ , \qquad t' = \cosh(\psi)t \ ,$$

therefore their ratio

$$\frac{x'_1}{t'} = \tanh(\psi) := v$$

gives the velocity of the new frame of reference with respect to the old. We note that $|v| = |\tanh(\psi)| < 1$. In terms of v, the hyperbolic rotation can be written in a more familiar form

$$H = \begin{pmatrix} \frac{1}{\sqrt{1-|v|^2}} & \frac{-v}{\sqrt{1-|v|^2}} \\ \frac{-v}{\sqrt{1-|v|^2}} & \frac{1}{\sqrt{1-|v|^2}} \end{pmatrix} .$$

The coordinate plane $\{(0, 0, x_2, \dots x_n)\}$ is invariant under $H(\psi)$, but the x_1-axis $\{(t, x_2, \dots x_n) = 0\}$ is moved;

$$x'_1 = \cosh(\psi)x_1 \ , \qquad t' = \sinh(\psi)x_1 \ ,$$

so that the $\{t = 0\}$ coordinate plane is tilted in space-time as well, at slope

$$\frac{x'_1}{t'} = \coth(\psi)$$

A diagram of the new space-time coordinates under a hyperbolic rotation is as follows.

The set that remains invariant under the transformations of the Lorentz group is the light cone itself. Setting $LC = \{t^2 - |x|^2 = 0\}$ this fact is checked as follows. Take $(t, x) \in LC$, then

$$\begin{pmatrix} t \\ x \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & -I_{n\times n} \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} = 0 .$$

Figure 1. This is the image of the $\{x = 0\}$ and $\{t = 0\}$ coordinate planes under the Lorentz transformation consisting of a hyperbolic rotation.

Under a Lorentz transformation $L : (t, x) \mapsto (t', x')$ we have

$$\begin{pmatrix} t' \\ x' \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & -I_{n\times n} \end{pmatrix} \begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} t \\ x \end{pmatrix}^T L^T \begin{pmatrix} 1 & 0 \\ 0 & -I_{n\times n} \end{pmatrix} L \begin{pmatrix} t \\ x \end{pmatrix} = 0 \ ,$$

and as we have checked above,

$$L^T \begin{pmatrix} 1 & 0 \\ 0 & -I_{n\times n} \end{pmatrix} L = \begin{pmatrix} 1 & 0 \\ 0 & -I_{n\times n} \end{pmatrix} \ .$$

The geometric interpretation of these properties is that the matrix

$$g := \begin{pmatrix} 1 & 0 \\ 0 & -I_{n\times n} \end{pmatrix}$$

defines the Minkowski metric $ds^2 = (dt, dx_1, \ldots dx_n)^T g(dt, dx_1, \ldots dx_n)$ on the space-time $\mathbb{R}^1_t \times \mathbb{R}^n_x$, which we have just shown to be invariant under the Lorentz group.

The invariance of the wave equation under the Lorentz group of transformations was part of a paradox that physics faced towards the end of the 19^{th} century. As we have seen in (1.1), Maxwell's equations are intimately tied to the wave equation. While the equations of classical mechanics are invariant under the Galilean transformation group, the theory of electricity and magnetism is invariant under the Lorentz group, and these are incompatible. Compatibility was restored in 1905 when Einstein introduced the special theory of relativity. However this was a revolutionary change in our perception of the universe, for which many intuitive ideas about space-time had to be modified. One of these we have seen above; the action of a hyperbolic rotation transforms the plane $\{t = 0\}$ of spatial coordinates as well as the time axis. Since we think of the spatial coordinate plane as being the 'present' state of the universe, it is a new idea that there is no universally valid instant that can globally be considered to be the present, and that the sense of the present is relative to the frame of reference of the observer. The concept of simultaneity has to be sacrificed. In terms of the wave equation, any one of the hypersurfaces $\{t' = 0\}$ can be used as a Cauchy surface for initial data for the wave equation.

6.3. Method of spherical means

There is another method for representing the solution of the wave equation in n space dimensions (6.1), based on the spherical means that we have encountered in Chapter 4, in formula (4.20). Recall the elementary solution method in the case of spatial dimension $n = 1$, via the d'Alembert formula

$$u(t,x) = \tfrac{1}{2}(f(x+t) + f(x-t)) + \tfrac{1}{2}\int_{x-t}^{x+t} g(y)\,dy \ .$$

The goal of this section is to produce similar expressions for the solution of the wave equation in the case of higher space dimensions.

Definition 6.4. Given a function $h(x) \in C(\mathbb{R}^n)$, its *spherical mean* centered about the point $x \in \mathbb{R}^n$ is

$$M(h)(x,r) := \frac{1}{\omega_n r^{n-1}} \int_{|x-y|=r} h(y)\,dS_y \ , \tag{6.7}$$

the average of $h(x)$ over a sphere $\mathbb{S}^{n-1}$ centered at x of radius r.

Recall that $\omega_n r^{n-1}$ is the surface area of the sphere $\mathbb{S}^{n-1}$ of radius r in n dimensions. When $n = 1$ the spherical mean of a function $f(x)$ is $M(f)(x,r) = \frac{1}{2}(f(x+r) + f(x-r))$, reminding one of one of the two terms of the d'Alembert formula. Changing variables in the integral (6.7), $y \mapsto x + r\xi$, $\xi \in S_1(0)$, there is another useful expression for the spherical mean;

$$M(h)(x,r) = \frac{1}{\omega_n} \int_{|\xi|=1} h(x+r\xi)\,dS_\xi \ .$$

This expression is defined for $r \geq 0$, but it is clear from this second expression that $M(h)(x,r)$ is an even function of r;

$$M(h)(x,-r) = M(h)(x,r) \ .$$

Lemma 6.5 (Darboux equation)**.** Given $h \in C^2(\mathbb{R}^n)$ then

$$\Delta_x M(h)(x,r) = \big(\partial_r^2 + \frac{n-1}{r}\partial_r\big) M(h)(x,r) \ ,$$

a formula that relates the Laplace operator of $M(h)$ in the n-dimensional x variables to its one dimensional r derivatives.

Proof. The Darboux equation follows from a calculation using multivariate calculus. First take one derivative;

$$\begin{aligned}\partial_r M(h)(x,r) &= \frac{1}{\omega_n} \int_{|\xi|=1} \partial_r h(x+r\xi)\,dS_\xi \\ &= \frac{1}{\omega_n} \int_{|\xi|=1} \sum_{j=1}^{n} \partial_{x_j} h(x+r\xi)\xi_j\,dS_\xi \ ,\end{aligned}$$

and we notice that $\sum_{j=1}^n \partial_{x_j} h(x + r\xi)\xi_j = \nabla h \cdot N$, the outwards normal derivative of h on the unit sphere. Continue this line of calculation using Green's theorem;

$$\begin{aligned}
\frac{1}{\omega_n}\int_{|\xi|=1}\sum_{j=1}^n \partial_{x_j} h(x+r\xi)\xi_j \, dS_\xi &= \frac{1}{\omega_n}\int_{|\xi|<1} r\Delta_x h(x+r\xi)\, d\xi \\
&= \frac{1}{\omega_n r^{n-1}}\Delta_x \int_{|x-y|<r} h(y)\, dy \\
&= \frac{1}{\omega_n r^{n-1}}\Delta_x \Big(\int_0^r \int_{|x-y|=\rho} h(y)\, dS_y \, d\rho\Big) \\
&= \frac{1}{r^{n-1}}\Delta_x \Big(\int_0^r \rho^{n-1} M(h)(x,\rho)\, d\rho\Big) \ .
\end{aligned}$$

The second step is to take a second derivative, after multiplying through by r^{n-1};

$$\begin{aligned}
\partial_r\big(r^{n-1}\partial_r M(h)(x,r)\big) &= \partial_r\Big(\Delta_x \int_0^r \rho^{n-1} M(h)(x,\rho)\, d\rho\Big) \\
&= \Delta_x\big(r^{n-1} M(h)(x,r)\big) \ .
\end{aligned}$$

Therefore

$$\begin{aligned}
\Delta_x M(h)(x,r) &= \frac{1}{r^{n-1}}\partial_r\big(r^{n-1}\partial_r M(h)(x,r)\big) \\
&= \big(\partial_r^2 + \frac{n-1}{r}\partial_r\big) M(h)(x,r) \ ,
\end{aligned} \tag{6.8}$$

which is the result of the lemma. □

Proposition 6.6. (i) For $h(x) \in C(\mathbb{R}^n)$ the value of $h(x)$ at any $x \in \mathbb{R}^n$ can be recovered from its spherical means;

$$h(x) = \lim_{r\to 0} M(h)(x,r) = M(h)(x,0) \ .$$

(ii) Additionally, for $h(x) \in C^2(\mathbb{R}^n)$

$$\partial_r M(h)(x,0) = \lim_{r\to 0} \frac{r}{\omega_n}\int_{|\xi|<1} h(x+r\xi)\, d\xi = 0 \ .$$

Formulae involving spherical means can be used to give an expression for the solution of the wave equation. Suppose that $u(t,x)$ is a solution of the Cauchy problems for the wave equation (6.1). Then its spherical mean $M(u)(t,x,r)$ satisfies an auxiliary equation in the reduced space-time variables $(t,r) \in \mathbb{R}^2$. Specifically, define

$$M(u)(t,x,r) = \frac{1}{\omega_n}\int_{|\xi|=1} u(t, x+r\xi)\, dS_\xi \ .$$

Taking time derivatives and using that $u(t,x)$ is a solution of (6.1), one obtains

$$\begin{aligned}\partial_t^2 M(u)(t,x,r) &= \frac{1}{\omega_n}\int_{|\xi|=1} \partial_t^2 u(t,x+r\xi)\, dS_\xi \\ &= \frac{1}{\omega_n}\int_{|\xi|=1} \Delta_x u(t,x+r\xi)\, dS_\xi = \Delta_x M(u)(t,x,r)\ .\end{aligned}$$

Now use the Darboux equation of Lemma 6.5

$$\partial_t^2 M(u)(t,x,r) = \big(\partial_r^2 + \frac{n-1}{r}\partial_r\big) M(u)(t,x,r)\ , \tag{6.9}$$

a partial differential equation in the two variables $(t,r) \in \mathbb{R}^2$. This is the *Euler – Poisson – Darboux* equation.

The wave equation in $\mathbb{R}^3$*:* The equation (6.9) is normally posed as an initial value problem

$$\begin{aligned}&\partial_t^2 M(u)(t,x,r) = \big(\partial_r^2 + \frac{n-1}{r}\partial_r\big) M(u)(t,x,r) \\ &M(u)(0,x,r) = M(f)(x,r)\ , \qquad \partial_t M(u)(0,x,r) = M(g)(x,r)\ .\end{aligned} \tag{6.10}$$

The precise methods and the character of the solution depend quite a bit on the the spatial dimension under consideration. The most straightforward case is in dimension $n = 3$. In the Euler – Poisson – Darboux equation (6.10) use the substitution $M(u)(t,x,r) \mapsto rM(u)(t,x,r)$, giving

$$\partial_t^2\big(rM(u)(t,x,r)\big) = r\big(\partial_r^2 + \frac{2}{r}\partial_r\big)M(u)(t,x,r) = \partial_r^2\big(rM(u)(t,x,r)\big)\ .$$

Therefore the function $v(t,r) = rM(u)(t,x,r)$ is a solution of the wave equation in one dimension, and the variable x is relegated to the role of a parameter. The solution is given by the d'Alembert formula

$$\begin{aligned}v(t,r) &= rM(u)(t,x,r) \\ &= \tfrac{1}{2}\big((r+t)M(f)(x,r+t) + (r-t)M(f)(x,r-t)\big) \\ &\quad + \tfrac{1}{2}\int_{r-t}^{r+t} \rho M(g)(x,\rho)\, d\rho\ .\end{aligned}$$

Since $M(f)(x,r)$ and $M(g)(x,r)$ are even functions under $r \mapsto -r$, then $rM(f)(x,r)$ and $rM(g)(x,r)$ are odd, therefore we may rewrite the above expression, dividing through by r;

$$\begin{aligned}M(u)(t,x,r) &= \frac{1}{2r}\big((t+r)M(f)(x,t+r) - (t-r)M(f)(x,t-r)\big) \\ &\quad + \frac{1}{2r}\int_{t-r}^{t+r} \rho M(g)(x,\rho)\, d\rho\ .\end{aligned}$$

We have used the fact that $\int_{t-r}^{r-t} \rho M(g)(x,\rho)\,d\rho = 0$ which holds because $rM(g)(x,r)$ is odd. Taking the limit as $r \to 0$, we recover a representation for the solution $u(t,x)$.

Theorem 6.7 (Kirchhoff's formula)**.** *When $n = 3$ the solution to the wave equation* (6.1) *is given by the expression*

$$u(t,x) = \partial_t\big(tM(f)(x,t)\big) + tM(g)(x,t) \ , \tag{6.11}$$

which is well defined as long as $f(x) \in C^1(\mathbb{R}^3)$ and $g(x) \in C(\mathbb{R}^3)$.

However it is quite explicit to see that in terms of pointwise regularity, the solution is in general less smooth than the initial data, for the Kirchhoff formula depends upon the the derivative of $f(x)$. Specifically, carrying out the differentiation in (6.11) we obtain

$$u(t,x) = \frac{1}{4\pi t^2}\int_{|x-y|=t} \big(tg(y) + f(y) + t\nabla f(y)\cdot \frac{x-y}{|x-y|}\big)\,dS_y \ .$$

Corollary 6.8. Given initial data $g(x) \in C^2(\mathbb{R}^3)$ and $f(x) \in C^3(\mathbb{R}^3)$ then the spherical means solution given in (6.11) is a *classica*l solution $u(t,x) \in C^2(\mathbb{R}^1_t \times \mathbb{R}^3_x)$.

This loss of differentiability for $n \geq 2$, due to focusing effects, is the topic of Problem 6.1. In contrast, when solutions are viewed in the sense of the Sobolev spaces H^s through the energy, there is no loss visible;

$$E(u)(t) = \int \tfrac{1}{2}(\partial_t u)^2 + \tfrac{1}{2}|\nabla_x u|^2\,dx = \int \tfrac{1}{2}g^2 + \tfrac{1}{2}|\nabla_x f|^2\,dx \ .$$

The Sobolev regularity of the solution is the same as the Sobolev regularity of the initial data.

The wave equation in $\mathbb{R}^n$ for odd dimensions n: There are similar expressions for the solution of the wave equation for $x \in \mathbb{R}^n$, for odd dimensions, for $n \geq 3$. Returning to the Euler – Poisson – Darboux equation (6.9) for the n dimensional wave equation

$$\partial_t^2 M(u) = \big(\partial_r^2 + \frac{n-1}{r}\partial_r\big)M(u) \ ,$$

we seek an algebraic reduction to the wave equation in two dimensions in the variables (t,r). This is able to be carried out in the odd dimensional case.

Proposition 6.9. Suppose that $k \geq 1$ is an integer and that $h = h(r) \in C^{k+1}(\mathbb{R}^1_+)$, then

$$\partial_r^2 \Big(\frac{1}{r}\partial_r\Big)^{k-1}\Big(r^{2k-1}h\Big) = \Big(\frac{1}{r}\partial_r\Big)^{k}\Big(r^{2k}\partial_r h\Big) \ ;$$
$$\Big(\frac{1}{r}\partial_r\Big)^{k-1}\Big(r^{2k-1}h\Big) = \sum_{j=0}^{k-1} \beta_j^k r^{j+1}\partial_r^j h \ .$$

with the combinatorial coefficients being $\beta_0^k = 1\cdot 3 \cdot \ldots (2k-1)) := (2k-1)!!$.

Proof. A proof by induction will work. □

The result is useful because we may set $n = 2k+1$ and take

$$v(t,x,r) := \Big(\frac{1}{r}\partial_r\Big)^{k-1}\Big(r^{2k-1}M(u)\Big)(t,x,r) \ ,$$

for a given solution $u(t,x)$ of (6.1). Then using the identities from Proposition 6.9

$$\begin{aligned}
\partial_r^2 v &= \partial_r^2 \Big(\frac{1}{r}\partial_r\Big)^{k-1}\Big(r^{2k-1}M(u)\Big) \\
&= \Big(\frac{1}{r}\partial_r\Big)^{k}\Big(r^{2k}\partial_r M(u)\Big) \\
&= \Big(\frac{1}{r}\partial_r\Big)^{k-1}\Big(r^{2k-1}\partial_r^2 M(u) + 2kr^{2k-2}\partial_r M(u)\Big) \\
&= \Big(\frac{1}{r}\partial_r\Big)^{k-1}\Big(r^{2k-1}\big(\partial_r^2 M(u) + \frac{2k}{r}\partial_r M(u)\big)\Big) \\
&= \Big(\frac{1}{r}\partial_r\Big)^{k-1}\Big(r^{2k-1}\partial_t^2 M(u)\Big) = \partial_t^2 v(t,r) \ .
\end{aligned}$$

Therefore $v(t,x,r)$ is the quantity that satisfies the one dimensional wave equation in the variables (t,r), for which we can apply the d'Alembert formula.

$$v(t,x,r) = \tfrac{1}{2}\big(v(0,r+t) + v(0,r-t)\big) + \tfrac{1}{2}\int_{r-t}^{r+t} \partial_t v(0,x,\rho)\, d\rho \ ,$$

with initial data

$$v(0,x,r) = \Big(\frac{1}{r}\partial_r\Big)^{k-1}\big(r^{2k-1}M(f)\big)(x,r) \ ,$$
$$\partial_t v(0,x,r) = \Big(\frac{1}{r}\partial_r\Big)^{k-1}\big(r^{2k-1}M(g)\big)(x,r) \ .$$

One recovers the solution as a limit

$$u(t,x) = \lim_{r\to 0} M(u)(t,x,r) = \lim_{r\to 0} \frac{v(t,x,r)}{\beta_0^k r} \tag{6.12}$$
$$= \frac{1}{(n-2)!!}\Big[\partial_t\Big(\frac{1}{t}\partial_t\Big)^{(n-3)/2}\big(t^{n-2}M(f)(t,x)\big) + \Big(\frac{1}{t}\partial_t\Big)^{(n-3)/2}\big(t^{n-2}M(g)(t,x)\big)\Big] .$$

Theorem 6.10. *For n odd, and for $f \in C^{(n-1)/2}(\mathbb{R}^n)$ and $g \in C^{(n-3)/2}(\mathbb{R}^n)$ the solution formula* (6.12) *gives a classical solution to the wave equation in $\mathbb{R}^n$.*

It is interesting to quantify the possible loss of smoothness of the solution over the initial data, made clear by the formula (6.12). Indeed the reduction process from $M(u)$ to v involves $k-1 = (n-3)/2$ derivatives, and the limit involves one derivative, therefore in general the solution will be less regular than the initial data by $k = (n-1)/2$ many derivatives.

6.4. Huygens' principle

Huygens' principle is the expression of the principle of finite propagation speed, analogous to the case of the one-dimensional wave equation in section 2. This general form of Huygens' principle is valid for the wave equation in any space dimension, and for hyperbolic equations in general. The strong form of Huygens' principle says more than this; it is the property that solutions are supported precisely on the union of light cones which have their vertex on the support of the initial data. There is a difference in dimension concerning this strong form of the Huygens' principle; it holds for odd space dimensional problems, but not in even dimensions. We state it here for the case of three space dimensions.

Theorem 6.11. *Consider solutions to the wave equation for $x \in \mathbb{R}^3$, and assume that the Cauchy data $f(x)$ and $g(x)$ are compactly supported, such that* $\mathrm{supp}(f) \cup \mathrm{supp}(g) \subseteq B_R(0)$*. Then*

(1) (Huygens' principle). *The solution $u(t,x)$ has its support within the bounded region $B_{R+|t|}(0)$;*

$$\mathrm{supp}(u(t,\cdot)) \subseteq B_{R+|t|}(0) .$$

(2) (strong Huygens' principle). *Additionally, for $|t| > R$, for any space-time point (t,x) in the region inside the light cones given by $\{(t,x) : |x| \le R-|t|\}$, again the solution vanishes; $u(t,x) = 0$.*

Proof. *(1). of Huygens' principle:* The result follows from the form of the Kirchhoff formula, which gives an expression for the solution $u(t,x)$ in terms of spherical means over spheres of radius $|t|$ in the initial hyperplane plane $\{t = 0\}$, which are of the form $\{y : |x - y| = |t|\}$ (the intersection of the backwards light cone emanating from (t, x) and the Cauchy hypersurface). If $|x| > R + |t|$ then this sphere does not intersect $B_R(0)$, which contains the support of the initial data, and hence the solution at that space-time point vanishes.

(2). of the strong Huygens' principle: When $|t| > R$ and $|x| < |t| - R$, again the backwards light cone emanating from the point (t, x) does not intersect the support of the initial data, in this case because the sphere $\{y : |x - y| = |t|\}$ is too big and has passed outside of $B_R(0)$.

One notes that this proof only depends upon the character of the solution representation as a spherical mean. Therefore the result holds for solutions of the wave equation in arbitrary odd dimensions, as exhibited in (6.12). □

In fact a more precise statement is true; at times $t > 0$ the solution $u(t, x)$ is supported within the region consisting of the union of light cones

$$LC_+(t, x) = \{(t, y) : t^2 - |y - x|^2 = 0, \, t > 0\}$$

whose vertices x lie in the set $\mathrm{supp}(f) \cup \mathrm{supp}(g)$.

A space-time picture of the support of the solution is as follows:

Figure 2. Space-time picture of the support

The finite propagation speed property is sometimes known as the weak Huygens' principle, it is a central feature of hyperbolic equations. It implies that in particular a signal will travel with finite speed. That is, for an observer standing still at point $x_0 \in \mathbb{R}^n$, $|x_0| > R$, the solution satisfies $u(x_0, t) = 0$ until $|x_0| - R < |t|$.

Corollary 6.12. If the initial data $f(x), g(x)$ vanish on $B_R(0)$, then the solution $u(t, x)$ must vanish on the cone $\{(t, x) \colon |x| < R - |t|\}$, $0 < |t| < R$. In particular if f, g are identically zero, then so will be $u(t, x)$.

Proof. This is a local uniqueness theorem. Suppose that $\mathrm{supp}(f) \cup \mathrm{supp}(g) \subseteq H_{a,v} = \{x \in \mathbb{R}^n \colon a \leq v \cdot x\}$ a half-space, then $\mathrm{supp}(u(t, x)) \subseteq H_{a+|t|,v}$. Therefore no data in any of the enveloping half-space $H_{R-|t|,v}, |v| = 1$ to $B_{R-|t|}(0)$ can propagate into the cone pictured below. □

Figure 3. The cone

Definition 6.13. (i) The *domain of influence* of a set $A \subseteq \mathbb{R}^n$ is the region in $\mathbb{R}^{n+1}$ in which a solution $u(x,t)$ of the wave equation can be affected by data in A.

(ii) The *domain of dependence* of a set $B \subseteq \mathbb{R}^{n+1}$ is the region $A \subseteq \mathbb{R}^n$ in which data can influence the solution in B.

The results of Theorem 6.17 and Corollary 6.12 imply the following pictures for the domains of influence and dependence of solutions of the wave equation.

Figure 4. Domain of influence of A

Figure 5. Domain of dependence of the point B

6.5. Paley-Wiener theory

It is self-evident that the complex exponentials $e^{-i\xi\cdot x}$, with $\xi \in \mathbb{R}^n$ and $x \in \mathbb{R}^n$, extend as holomorphic functions $e^{-i\zeta\cdot x}$ for $\zeta \in \mathbb{C}^n$. More specifically to its character, the extension is defined over all of $\mathbb{C}^n$, meaning that is it is *entire*, and furthermore the extension has bounds on its growth at infinity of the form

$$|e^{-i\zeta\cdot x}| \le e^{|x||im(\zeta)|}.$$

Definition 6.14. An entire function $g(\zeta)$, $\zeta \in \mathbb{C}^n$, is of *exponential type R* if it satisfies the estimates

$$(1+|\zeta|^2)^{N/2}|g(\zeta)| \le C_N e^{R|im(\zeta)|} \tag{6.13}$$

for all $N \in \mathbb{N}$.

Examples of the behavior are given by the Fourier transform. The complex exponential above is not strictly speaking of exponential type because it only satisfies (6.13) for the case $N = 0$. Another example of an entire function, which however is not of exponential type, is

$$g(\zeta) = e^{-\zeta^2/2} = e^{-\frac{1}{2}(\xi^2-\eta^2)-i\xi\eta} ,$$

for $\zeta = \xi + i\eta \in \mathbb{C}$. Given any function $f \in \mathcal{S}$, we can express it's Fourier transform as a superposition of complex exponentials:

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} e^{-i\xi\cdot x} f(x)\, dx \ ;$$

where each complex exponential extends to an entire function on $\mathbb{C}^n$ but the function $\hat{f}(\xi)$ itself need not necessarily have a holomorphic extension off of the real subspace $\mathbb{R}^n \subseteq \mathbb{C}^n$ at all. However if a function $g(\xi)$ is such that $g(\xi) = \hat{f}(\xi)$ with $f(x) \in C_0^\infty(\mathbb{R}^n)$, in particular if $f(x)$ has compact support, then $g(\xi)$ does have an entire holomorphic extension.

Proposition 6.15. Suppose that $f \in C_0^\infty(\mathbb{R}^n)$, with $\mathrm{supp}(f) \subseteq \{x \in \mathbb{R}^n \colon |x| \leq R\}$. Then $\hat{f}(\xi) = g(\xi)$ extends to an entire function, which is of exponential type R.

Proof. For each $\zeta \in \mathbb{C}^n$, the integral

$$g(\zeta) = \frac{1}{\sqrt{2\pi}^n} \int_{B_R(0)} e^{-i\zeta\cdot x} f(x) dx$$

converges absolutely, uniformly over bounded sets of ζ. To check that $g(\zeta)$ is holomorphic we simply test the Cauchy-Riemann equations:

$$\partial_{\bar{\zeta}} g(\zeta) = \frac{1}{\sqrt{2}}(\partial_\xi + i\partial_\eta) g(\zeta) = \frac{1}{\sqrt{2\pi}^n} \int \partial_{\bar{\zeta}}(e^{-i\zeta\cdot x}) f(x)\, dx = 0\ .$$

Lastly, for α a multi-index with $|\alpha| = N$, we need to estimate the integrals

$$\left| \zeta^\alpha \int_{\mathbb{R}^n} e^{-i\zeta\cdot x} f(x) dx \right| = \left| \int_{B_R(0)} e^{-i\zeta\cdot x} \left(\frac{1}{i}\partial_x\right)^\alpha f(x) dx \right|$$
$$\leq \int_{B_R(0)} e^{|x||im(\zeta)|} |\partial_x^\alpha f(x)| dx \leq C_{|\alpha|} e^{R|im(\zeta)|}.$$

□

The content of Paley-Wiener theory is to say that the converse also holds.

Theorem 6.16 (Paley-Wiener)**.** *Suppose that $g(\zeta)$ is an entire function of exponential type R. Then there is a function $f(x) \in C_0^\infty(\mathbb{R}^n)$ with $\mathrm{supp}(f) \subseteq B_R(0)$ such that*

$$g(\zeta) = \hat{f}(\zeta) = \frac{1}{\sqrt{2\pi}^n} \int e^{-i\zeta\cdot x} f(x)\, dx\ .$$

Proof. It is straightforward to restrict $g(\xi)$ to $\xi \in \mathbb{R}^n$, and define

$$f(x) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} e^{i\xi\cdot x} g(\xi) d\xi.$$

Since $|g(\xi)| \leq C_N(1 + |\xi|^2)^{-N/2}$ as in (6.13), this integral is absolutely convergent. We may also take an arbitrary number of derivatives of $f(x)$

$$\partial_x^\beta f(x) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} e^{i\xi\cdot x} (-i\xi)^\beta g(\xi)\, d\xi\ ,$$

and the integral remains absolutely convergent. Thus $f \in C^\infty(\mathbb{R}^n)$ and the question that remains has to do with its support.

Using that $g(\zeta)$ is holomorphic, consider deformations of the region of integration off of the real axis in the complex space $\mathbb{C}^n$,

$$\frac{1}{\sqrt{2\pi}^n}\int e^{i((\xi_1+i\eta_1)x_1+\xi'\cdot x')}g(\xi_1+i\eta_1,\xi')\,d\xi_1 d\xi'\ ,$$

for $\eta_1\in\mathbb{R}$. This is independent of η_1, as can be shown by Cauchy's theorem, taking the limit of an integral in $\zeta_1=\xi_1+i\eta_1$ over the countour and letting

Figure 6. The contour

$T\to+\infty$. The decay condition (6.13) assumes that there are no contributions from the boundaries $\xi_1=\pm T$ in the limit. Repeating this argument in all variables, we show that for any desired $\eta\in\mathbb{R}^n$,

$$f(x)=\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} e^{i(\xi+i\eta)\cdot x}g(\xi+i\eta)\,d\xi\ .$$

Now fix $x\neq 0$ in $\mathbb{R}^n$, and choose the particular $\eta=\lambda\frac{x}{|x|}$, with a real parameter $\lambda>0$. Then

$$\begin{aligned}|f(x)|\leq&\frac{1}{\sqrt{2\pi}^n}\left|\int_{\mathbb{R}^n} e^{i\xi\cdot x-\lambda|x|}g\left(\xi+i\lambda\frac{x}{|x|}\right)d\xi\right|\\ \leq&\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} e^{-\lambda|x|}\left|g\left(\xi+i\lambda\frac{x}{|x|}\right)\right|d\xi\\ \leq&C_N\int_{\mathbb{R}^n}(1+|\xi|^2)^{-N/2}e^{-\lambda|x|+\lambda R}d\xi.\end{aligned}$$

Now suppose that $|x|>R$; as $\lambda\to+\infty$ the RHS tends to zero. Thus we have shown that $f(x)=0$. This proves that indeed $\mathrm{supp}(f)\subseteq B_R(0)$. □

There is a similar theory for distributions of compact support, and their Fourier transforms as entire functions of exponential type; conversely if $g(\zeta)$ is an entire function which for some N satisfies the estimate

$$|g(\zeta)|\leq C(1+|\zeta|^2)^{N/2}e^{R|im(\zeta)|}, \tag{6.14}$$

then there is a distribution $f\in\mathcal{D}'$, with $\mathrm{supp}(f)\subseteq B_R(0)$ such that the (generalized) Fourier transform of f is g.

Hugens principle revisited: Returning to the discussion of solutions of the wave equation, our Fourier integral expression is that

$$u(x,t)=\frac{1}{\sqrt{2\pi}^n}\int e^{i\xi\cdot x}\left(\cos(|\xi|t)\hat{f}(\xi)+\frac{\sin(|\xi|t)}{|\xi|}\hat{g}(\xi)\right)d\xi\ .$$

Let us suppose that the initial data is of compact support:

$$\mathrm{supp}(f)\cup\mathrm{supp}(g)\subseteq B_R(0)=\{x\in\mathbb{R}^n\colon |x|<R\}\ .$$

Then by Proposition 6.15, both $\hat{f}(\xi)$ and $\hat{g}(\xi)$ extend to entire functions on $\mathbb{C}^n$, of exponential type R. Let us examine the Fourier transform of the solution

$$\hat{u}(\xi,t) = \cos(|\xi|t)\hat{f}(\xi) + \frac{\sin(|\xi|t)}{|\xi|}\hat{g}(\xi) \ .$$

The individual functions $\cos(\sqrt{\zeta^2}t)\hat{f}(\zeta)$ and $\frac{\sin(\sqrt{\zeta^2}t)}{\sqrt{\zeta^2}}\hat{g}(\zeta)$ are entire functions of $\zeta \in \mathbb{C}^n$, where we are using the notation that $\zeta^2 = \zeta_1^2 + \zeta_2^2 + \ldots + \zeta_n^2$, for $\zeta_j \in \mathbb{C}$. Furthermore,

$$\left|\cos(\sqrt{\zeta^2}t)\right| \le e^{|im(\zeta)||t|}, \qquad \left|\frac{\sin(\sqrt{\zeta^2}t)}{\sqrt{\zeta^2}}\right| \le e^{|im(\zeta)||t|},$$

therefore the two products $\cos(\sqrt{\zeta^2}t)\hat{f}(\zeta)$ and $\left(\sin(\sqrt{\zeta^2}t)/\sqrt{\zeta^2}\right)\hat{g}(\zeta)$ are of exponential type $(R+|t|)$. We can conclude that the solution $u(x,t)$ of the wave equation is the Fourier transform of an entire function of exponential type $(R+|t|)$, and thus by the Paley –Wiener theorem the solution $u(x,t)$ has its support in the ball $B_{R+|t|}(0)$ of radius $R+|t|$.

Theorem 6.17 (Huygens' principle again)**.** *Solutions of the wave equation* (6.1) *which start with initial data with support satisfying* $\mathrm{supp}(f) \cup \mathrm{supp}(g) \subseteq B_R(0)$ *satisfy at nonzero times* $t \in \mathbb{R}$

$$\mathrm{supp}(u(t,x)) \subseteq B_{R+|t|}(0).$$

6.6. Lagrangians and Hamiltonian PDEs

The wave equation can be derived from a Lagrangian, using a compelling analogy with classical mechanics. The *Lagrangian* function in the case of the wave equation (6.1) is defined by

$$L := \int_{\mathbb{R}^n} \tfrac{1}{2}(\partial_t u)^2 - \tfrac{1}{2}|\nabla u|^2 \, dx \ . \tag{6.15}$$

This Lagrangian has an associated *action integral* given by

$$S = \int_0^T \int_{\mathbb{R}^n} \tfrac{1}{2}(\partial_t u)^2 - \tfrac{1}{2}|\nabla u|^2 \, dxdt \ . \tag{6.16}$$

From the action the wave equation arises from the *principle of least action*, which dictates that the motion of a Lagrangian system is a stationary point of the action integral. A stationary point $u(t,x)$ of the action over the time interval $0 \le t \le T$ satisfies $\delta S = 0$ for all admissible variations $v(t,x) = \delta u(t,x)$, where $\delta u(t,x)$ denotes a small but arbitrary variation of the function $u(t,x)$. Admissible variations are smooth, and are such that $v(0,x) = v(T,x) = 0$, so that $u(t,x)$ and $u(t,x) + v(t,x)$ have the same initial and final states over the time interval $[0,T]$. In the case of the wave equation, a stationary point of the action satisfies

$$\begin{aligned}\delta S &= \int_0^T \int_{\mathbb{R}^n} \partial_t u \partial_t v - \nabla_x u \cdot \nabla_x v \, dxdt \\ &= -\int_0^T \int_{\mathbb{R}^n} (\partial_t^2 u - \Delta u) v \, dxdt + \int_{\mathbb{R}^n} \partial_t u v \, dx \Big|_{t=0}^T .\end{aligned}$$

From the assumption that $v(0,x) = v(T,x) = 0$, the last term of the RHS vanishes. The conclusion is that, because $v(t,x)$ is otherwise arbitrary, we must have that

$$\partial_t^2 u - \Delta u = 0 .$$

These are the *Euler – Lagrange* equations for the action (6.16). Notice that the principle of stationary action allows us to give an initial *position* $u(0,x) = f(x)$ but does not allow for setting the initial *momentum* $\partial_t u(0,x) = g(x)$, so it is not in fact compatible with the initial value problem. Nonetheless the principle of least action, or more generally of stationary action, remains a guiding principle for many equations in physics, while on a rigorous mathematical level the principle remains a formal one in this and other cases.

A Lagrangian L and subsequently an action integral S can be defined for more general systems, indeed this is how wave equations are derived in most problems in physics. Consider the more general Lagrangian

$$L = \int_{\mathbb{R}^n} \tfrac{1}{2}\dot{u}^2 - G(u, \nabla_x u) \, dx \tag{6.17}$$

whose associated action is given by

$$S = \int_0^T \int_{\mathbb{R}^n} \tfrac{1}{2}\dot{u}^2 - G(u, \nabla_x u) \, dxdt .$$

Suppose that the field $u(t,x)$ is a stationary point of the action;

$$\begin{aligned}0 = \delta S &= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \int_0^T \int_{\mathbb{R}^n} \tfrac{1}{2}(\partial_t u + \varepsilon \partial_t v)^2 - G(u + \varepsilon v, \nabla_x (u + \varepsilon v)) \, dxdt \\ &= \int_0^T \int_{\mathbb{R}^n} \partial_t u \partial_t v - \partial_{\nabla u} G(u, \nabla_x u) \cdot \nabla_x v - \partial_u G(u, \nabla_x u) v \, dxdt \\ &= \int_0^T \int_{\mathbb{R}^n} \big[-\partial_t^2 u + \nabla_x \cdot \partial_{\nabla u} G(u, \nabla_x u) - \partial_u G(u, \nabla_x u)\big] v \, dxdt \\ &\quad + \int_{\mathbb{R}^n} \partial_t u v \, dx \Big|_{t=0}^T .\end{aligned}$$

The notation is that the nonlinear function $G = G(u,V)$ depends upon the variables u as well as the n components of $V = \nabla u$. The notation for its partial derivatives is that $\partial_{\nabla u} G(u, \nabla_x u) = \partial_V G(u,V)|_{V=\nabla_x u}$. The final term vanishes because $v(t,x)$ is an admissible variation. Since v is otherwise arbitrary, the field $u(t,x)$ must satisfy the Euler – Lagrange equations

$$\partial_t^2 u - \nabla_x \cdot \partial_{\nabla u} G(u, \nabla_x u) + \partial_u G(u, \nabla_x u) = 0 \ .$$

This is a hyperbolic equation if the matrix of partial derivatives of G with respect to the variables V is positive definite. In the example of the wave equation, $G(u, \nabla u) = \frac{1}{2}|\nabla u|^2$, and $\partial^2_{\nabla u} G = I$.

There is a compelling analogy with classical mechanics in this formal treatment of field theories. To continue the analogy, given a Lagrangian $L(\dot{u}, u)$, there is a transformation of the Euler – Lagrange equations to a Hamiltonian system, in our case to a system of Hamiltonian PDEs. This is illustrated in the example Lagrangian

$$L = \int_{\mathbb{R}^n} \tfrac{1}{2}(\dot{u})^2 - G(\nabla_x u)\, dx \ , \tag{6.18}$$

where for clarity we have simplified the Lagrangian (6.17) above. The action integral is as before

$$S = \int_0^T L\, dt,$$

and as above the principle of stationary action gives the Euler – Lagrange equations

$$-\partial_t \delta_{\dot{u}} L + \delta_u L = 0 \ , \tag{6.19}$$

where $\delta_u L = \nabla_x \cdot \partial_{\nabla u} G(\nabla_x u)$ and $\delta_{\dot{u}} L = \dot{u}$. In general a Lagrangian may depend explicitly on time $L(\dot{u}, u, t)$, but in many cases, such as the one at hand, it describes a physical process whose properties do not change with time, and it does not. In this situation, the Euler – Lagrange equations exhibit a conservation law. This can be seen from the following computation:

$$\begin{aligned} \frac{d}{dt} L &= \int \big(\delta_{\dot{u}} L\, \ddot{u} + \delta_u L\, \dot{u}\big)\, dx = \int \big(\delta_{\dot{u}} L\, \ddot{u} + \partial_t(\delta_{\dot{u}} L)\, \dot{u}\big)\, dx \\ &= \frac{d}{dt} \int \delta_{\dot{u}} L\, \dot{u}\, dx \ . \end{aligned}$$

We used the Euler – Lagrange equations (6.19) in the last equality of the first line. Therefore the conservation law is evident, namely

$$\frac{d}{dt}\Big(\int \delta_{\dot{u}} L\, \dot{u} - L\Big) = 0 \ . \tag{6.20}$$

In the example (6.18) this conservation law is

$$\frac{d}{dt} \int_{\mathbb{R}^n} \tfrac{1}{2}(\partial_t u)^2 + G(\nabla_x u)\, dx = 0 \ .$$

Definition 6.18. (i) Define the *Hamiltonian* of the system (6.19) by

$$H := \int \delta_{\dot{u}} L \, \dot{u} \, dx - L$$

(ii) The *conjugate momentum* of the field u is defined to be

$$p := \delta_{\dot{u}} L \ ,$$

giving $p = p(\dot{u}, u)$.
(iii) If the relationship between $\dot{u}$ and p can be inverted to obtain $\dot{u} = \dot{u}(p, u)$, at least locally, then the mapping $\dot{u} \to p$ is called the *Legendre transform.* Using this mapping, we may rewrite the Hamiltonian

$$H = \int \dot{u} \, p \, dx - L = H(u, p) \ ,$$

in terms of the new Hamiltonian variables (u, p).

The Legendre transform offers an elegant way to transform a second order equation (in time) to a first order system of equations.

Theorem 6.19. *The Euler – Lagrange equations* (6.19) *for* $u = u(t, x)$ *are equivalent in the new variables* $(u(t, x), p(t, x))$ *to the system*

$$\begin{aligned} \partial_t u &= \delta_p H \\ \partial_t p &= -\delta_u H \ . \end{aligned} \tag{6.21}$$

The system of equation (6.21) is known as *Hamilton's canonical equations* for the evolution equations described by H.

Proof. The formal equivalence of (6.19) and (6.21) is a general fact. Firstly,

$$H = \int \dot{u} \, p \, dx - L \ ,$$

so that $\dot{u} = \delta_p H$. Secondly we note that $\delta_u H = -\delta_u L$, so that

$$\dot{p} = \frac{d}{dt}\big(\delta_{\dot{u}} L\big) = \delta_u L = -\delta_u H \ .$$

□

Exhibiting this transformation in the setting of the wave equation, we have

$$L = \int_{\mathbb{R}^n} \tfrac{1}{2}\dot{u}^2 - \tfrac{1}{2}|\nabla_x u|^2 \, dx \ ,$$

from which the Legendre transform gives $p = \delta_{\dot{u}} L = \dot{u}$. Then

$$H = \int \dot{u} \, p \, dx - L(u, \dot{u}) = \int \dot{u}^2 \, dx - L = \int \tfrac{1}{2}p^2 + \tfrac{1}{2}|\nabla_x u|^2 \, dx \ .$$

Hamilton's canonical equations are then

$$\begin{aligned} \partial_t u &= \delta_p H = p \\ \partial_t p &= -\delta_u H = \Delta u \ , \end{aligned} \tag{6.22}$$

which is of course the wave equation presented as a first order system of equations. The energy functional $E(u)$ for the wave equation is the Hamiltonian $H(u,p)$ for the system.

Hamiltonian PDEs: This formal exposition is a lead-in to study other PDEs that can be posed in the form of Hamiltonian systems. Consider systems of equations of the form

$$\partial_t z = J \mathrm{grad}_z H \ , \tag{6.23}$$

where z is a vector function, and the matrix $J = -J^T$ is skew symmetric. Rewriting (6.22) as

$$\partial_t \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} p \\ \Delta u \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} -\Delta u \\ p \end{pmatrix} := J \mathrm{grad}_{(u,p)} H \ , \tag{6.24}$$

the wave equation takes this form, with

$$J := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \ ,$$

where we note that the matrix J satisfies $J = -J^T$.

Proposition 6.20 (Conservation of energy)**.** The Hamiltonian $H(u,p)$ is a conerved quantity for solutions of (6.24).

Proof. This can reasonably be called the law of conservation of energy as the Hamiltonian function is often, not always, the energy of the system being considered. The following classical calculation using the chain rule makes the assumption that solutions to (6.23) exist. Indeed from (6.23),

$$\frac{d}{dt} H(z(t)) = \langle \mathrm{grad}_z H, \dot{z} \rangle = \langle \mathrm{grad}_z H, J \mathrm{grad}_z H \rangle = 0 \ ,$$

where we have used the skew symmetry of the matrix J. □

A number of other problems can be viewed as Hamiltonian systems in infinitely many variables. Several examples are:

1. nonlinear wave equations. Hyperbolic equations of this form can be treated as Hamiltonian PDEs. Define the Hamiltonian to be

$$H(u,p) = \int_{\mathbb{R}^n} \tfrac{1}{2} p^2 + \tfrac{1}{2} |\nabla u|^2 + G(x,u) \, dx$$

then the gradient of H is given by

$$\delta H(u,p) = \begin{pmatrix} -\Delta u + \partial_u G(x,u) \\ p \end{pmatrix}$$

and therefore the equations of motion are posed as

$$\partial_t \begin{pmatrix} u \\ p \end{pmatrix} = J \mathrm{grad} H \ .$$

which is a first order system equivalent to the nonlinear wave equation

$$\partial_t^2 u = \Delta u - \partial_u G(x, u) \ .$$

2. nonlinear Schrödinger equations. The Hamiltonian for this set of equations is in the form

$$H(\psi, \bar{\psi}) = \int \tfrac{1}{2} |\nabla \psi|^2 + Q(x, \psi, \bar{\psi}) \, dx \ ,$$

where $Q(x, \psi, \xi) : \mathbb{R}_x \times \mathbb{C}_\psi \times \mathbb{C}_\xi \to \mathbb{C}$ has the property that $Q(x, \psi, \bar{\psi})$ is real valued. Then $\delta_{\bar{\psi}} H(\psi, \bar{\psi}) = -\frac{1}{2} \Delta_x \psi + \partial_{\bar{\psi}} Q$, and setting $J = iI$ (which is a nondegenerate, skew symmetric operator), then

$$\begin{aligned} \partial_t \psi &= J \delta_{\bar{\psi}} H \\ &= i(-\frac{1}{2} \Delta \psi + \partial_{\bar{\psi}} Q(x, \psi, \bar{\psi})) \ . \end{aligned}$$

When $Q = \pm |\psi|^4$ this equation is the well-known cubic nonlinear Schrödinger equation, where the $+$ sign is the defocusing case and the $-$ sign is the focusing case.

3. Korteweg-de Vries equation (KdV). This famous dispersive equation first arose in the 19^{th} century as a model of waves in the free surface of water in a canal. Currently it is used in modeling numerous phenomena including tsunami propagation. It is also well known as a PDE that is a completely integrable Hamiltonian system, where the study of the phase space of solutions has led to discoveries as far ranging as algebraic geometry and inverse spectral theory. The Hamiltonian is

$$H(q) = \int_{-\infty}^{\infty} \frac{1}{12} (\partial_x q)^2 + G(q) \, dx \ .$$

As above, the gradient of $H(q)$ is given by

$$\delta_q H = -\frac{1}{6} \partial_x^2 q + \partial_q G(q) \ .$$

Setting $J = \partial_x$, which again is a skew symmetric operator, we arrive at Hamilton's canonical equations in the form

$$\begin{aligned} \partial_t q &= \partial_x (-\frac{1}{6} \partial_x^2 q + \partial_q G(q)) \\ &= -\frac{1}{6} \partial_x^3 q + G''(q) \partial_x q \ . \end{aligned}$$

The most well-known versions of the KdV equation are when $G(q) = \frac{1}{3}q^3$, and when $G(q) = \frac{1}{4}q^4$.

Exercises: Chapter 6

Exercise 6.1. (Focusing singularity of solutions of the wave equation in $\mathbb{R}^3$ (d'après F. John)):

(i) Suppose that the initial data for the wave equation in three dimensions has spherically symmetric data;

$$f(x) = f(r) \ , \quad g(x) = g(r) \ , \qquad r^2 = x_1^2 + x_2^2 + x_3^2 \ .$$

Show that the general solution can be expressed as

$$u(t,r) = \frac{1}{r}\big(F(r+t) + G(r-t)\big) \ ,$$

that is, it consists of an incoming wave and an outgoing wave.

(ii) With the special initial data $u(0,r) = 0$, $\partial_t u(r) = g(r)$ with $g(r)$ an even function of r, then

$$u(t,r) = \frac{1}{2r}\int_{r-t}^{r+t} \rho g(\rho)\, d\rho \ .$$

(iii) Set the initial data to be

$$g(r) = 1 \ , \quad 0 \le r < 1 \ , \quad g(r) = 0 \ , 1 \le r \ ,$$

show that $u(t,r)$ is continuous for $|t| < 1$ but at time $t = 1$ it exhibits a jump discontinuity. This is due to the focusing of the singularity in $\partial_t u(0,r)$ given at $t = 0$.

Exercise 6.2. This problem addresses the decay rate of solutions of the wave equation in $\mathbb{R}^3$. Suppose that the initial data $(f(x), g(x)) \in C^1(\mathbb{R}^3) \times C(\mathbb{R}^3)$ and that it is supported in the bounded set $B_1(0)$. By inspecting the Kirchhoff formula for the solution $u(t,x)$, show that

$$|u(t,x)| \le \frac{C}{|t|}$$

for some constant C, which can be quantified using $\|f\|_{C^1(\mathbb{R}^3)}$ and $\|g\|_{C(\mathbb{R}^3)}$.

Exercise 6.3. (Global existence with small initial data for certain nonlinear wave equations):

This question is to show that certain nonlinear wave equations possess smooth solutions for all $t \in \mathbb{R}$. This contrasts with other cases where solutions form singularities in finite time. Consider the equation

(6.25)
$$\partial_t^2 v - \Delta v + (\partial_t v)^2 - |\nabla v|^2 = 0$$
$$v(0,x) = f(x) \quad \partial_t v(0,x) = g(x) \qquad (f(x), g(x)) \in C^1(\mathbb{R}^3) \times C(\mathbb{R}^3) \ ,$$

with f, g supported in a compact set.

(i) Setting $u = e^v - 1$, show that $u(t,x)$ satisfies the wave equation

$$\partial_t^2 u - \Delta u = 0$$
$$u(0,x) = e^{f(x)} - 1 := F(x) \quad \partial_t u(0,x) = g(x) e^{f(x)} := G(x)$$
$$(F(x), G(x)) \in C^1(\mathbb{R}^3) \times C(\mathbb{R}^3) \ .$$

Explain why (F, G) has compact suport.

(ii) Show that for sufficiently small $\|(F, \partial_x F, G)\|_C$, the solution $u(t,x)$ is bounded by

$$|u(t,x)| < 1 \ ,$$

using the result of Problem 2.

In this case the transformation $v \mapsto u$ is invertible for all $(t,x) \in \mathbb{R}_t^1 \times \mathbb{R}_x^3$, giving rise to a global solution $v(t,x)$ of the equation (6.25).

Exercise 6.4. (method of decent for the wave equation for $x \in \mathbb{R}^2$)

(i) Show that if $x \in \mathbb{R}^3$ but the Cauchy data for the wave equation only depends upon (x_1, x_2), namely

$$f = f(x_1, x_2) \ , \qquad g = g(x_1, x_2) \ , \tag{6.26}$$

then the solution of the wave equation in $\mathbb{R}_t^1 \times \mathbb{R}_x^3$ is also independent of x_3;

$$u = u(t, x_1, x_2) \ ,$$

and therefore $u(t, x_1, x_2)$ satisfies the wave equation in two space dimensions;

$$\partial_t^2 u - (\partial_{x_1}^2 + \partial_{x_2}^2) u = 0 \ .$$

(ii) Use the Kirchhoff formula to express the solution to the wave equation in $\mathbb{R}^3$ for data satisfying (6.26).

(iii) In the expression in (ii) reparametrize the spherical integrals by their projection onto the (x_1, x_2)-plane; *e.g.*

$$\int_{\mathbb{S}^2 : |x-y|=t} f(y)\, dS_y = \iint_{|(x_1-y_1, x_2-y_2)|<t} f(y_1, y_2) \sqrt{t^2 - ((x_1-y_1)^2 + (x_2-y_2)^2)}\, dy_1 dy_2 \ ,$$

which gives a general formula in $\mathbb{R}^2$ for the solution of the wave equation

$$\Box u = 0 \ .$$

(iv) Describe the nature of this solution in the case that the support of f and g as functions on $\mathbb{R}^2$ is compact, say supported in the ball $B_R(0) = \{|(x_1, x_2)| < R\}$. In particular comment on the Huygens' principle. Does the solution satisfy the strong form of Huygens' principle, and why? Describe what an observer sees as time progresses when they are situated farther than R from the origin.