First order scalar semilinear equations have the form
\begin{equation}
\label{eq:1}
a(x, y)u_x + b(x, y)u_y = c(x, y, u);
\end{equation}
here we assume that \(a, b, c\) are at least \(C^1\), given real valued functions. Let \(V\) be the vector field on \(\mathbb{R}^2\) given by
\[
V(x, y) = (a(x, y), b(x, y)),
\]
so \(a(x, y)u_x + b(x, y)u_y\) is the directional derivative of \(u\) along \(V\). Let \(\gamma = \gamma(s)\) be an integral curve of \(V\), i.e. \(\gamma(s) = (x(s), y(s))\) has tangent vector \(V = V(x(s), y(s))\) for each \(s\). Explicitly, this says that
\begin{equation}
\label{eq:2}
x'(s) = a(x(s), y(s)), \quad y'(s) = b(x(s), y(s)).
\end{equation}
Now let \(v(s) = u(\gamma(s)) = u(x(s), y(s))\). Thus, by the chain rule,
\[
v'(s) = x'(s)u_x(x(s), y(s)) + y'(s)u_y(x(s), y(s))
\begin{align*}
&= a(x(s), y(s))u_x(x(s), y(s)) + b(x(s), y(s))u_y(x(s), y(s)) \\
&= c(x(s), y(s), u(x(s), y(s))),
\end{align*}
where in the last step we used the PDE. Thus,
\[
v'(s) = c(x(s), y(s), v(s)),
\]
i.e. \(v\) satisfies an ODE along each integral curve of \(V\).

To solve the PDE, we parameterize the integral curves by an additional parameter \(r\), i.e. the integral curves are \(\gamma_r = \gamma_r(s) = (x_r(s), y_r(s))\), where \(r\) is in an interval (or the whole real line), and each \(\gamma_r\) is an integral curve for \(V\), i.e.
\begin{equation}
\label{eq:3}
x'_r(s) = a(x_r(s), y_r(s)), \quad y'_r(s) = b(x_r(s), y_r(s)),
\end{equation}
so \(v_r(s) = u(\gamma_r(s))\) solves
\begin{equation}
\label{eq:4}
v'_r(s) = c(x_r(s), y_r(s), v_r(s)).
\end{equation}
Note that here the subscript \(r\) denotes a parameter, not a derivative! We may equally well write \(x_r(s) = x(r, s), \ y_r(s) = y(r, s)\), and we will do so; we adopted the subscript notation to emphasize that along each integral curve \(r\) is fixed, i.e. is a constant.

Which parameterization should we use? We are normally also given initial conditions along a curve \(\Gamma = \Gamma(r)\) with \(\Gamma(r) = (\Gamma_1(r), \Gamma_2(r))\), namely
\[
u(\Gamma(r)) = \phi(r),
\]
where \(\phi\) is a given function. For example, we are given an initial condition on the \(x\) axis: \(u(x, 0) = \phi(x)\), in which case we may choose \(r = x, \ \Gamma(r) = (r, 0)\). Then we want the integral curve with parameter \(r\) to go through \(\Gamma(r)\) at ‘time’ 0, i.e. we want
\[
\gamma_r(0) = \Gamma(r),
\]
and we want
\[
v_r(0) = u(\gamma_r(0)) = \phi(r).
\]
Combining these, we have two groups of ODEs: a system for the integral curves of \( V \), also called \textit{(projected) characteristic curves}, with initial conditions given by \( \Gamma \), and a scalar ODE along the integral curves with initial condition given by \( \phi \):

\[
\begin{align*}
x'_r(s) &= a(x_r(s), y_r(s)), \quad x_r(0) = \Gamma_1(r), \\
y'_r(s) &= b(x_r(s), y_r(s)), \quad y_r(0) = \Gamma_2(r),
\end{align*}
\]  
\tag{5}

and

\[
\begin{align*}
v'_r(s) &= c(x_r(s), y_r(s), v_r(s)), \quad v_r(0) = \phi(r).
\end{align*}
\]  
\tag{6}

We solve these by first solving the ODE in (5) with the initial conditions, then solving the ODE in (6) with the initial conditions. Then finally we express \((r, s)\) in terms of \(x, y\), i.e. we invert the map 

\[(r, s) \mapsto (x(r, s), y(r, s)), \]

to get \(r = R(x, y), \ s = S(x, y)\), and then our solution is

\[u(x, y) = v_{R(x, y)}(S(x, y)).\]

Let’s do concrete examples, starting with a simple (constant coefficient) homogeneous linear PDE.

\textbf{Problem 1.} Solve \( au_x + bu_y = 0\), where \(a, b\) are constants, \(a \neq 0\), with the initial condition \(u(0, y) = e^y\).

\textit{Solution.} Our initial curve will be \(\Gamma(r) = (0, r)\), and as \(y = r\) along \(\Gamma\), the initial condition is \(u(\Gamma(r)) = e^r\). The equations for the (projected) characteristic curves are

\[
\begin{align*}
x'_r(s) &= a, \quad x_r(0) = 0, \\
y'_r(s) &= b, \quad y_r(0) = r.
\end{align*}
\]

The solution is

\[
x_r(s) = as, \quad y_r(s) = r + bs.
\]

The ODEs along the characteristic curves are

\[
v'_r(s) = 0, \quad v_r(0) = e^r.
\]

The solution is \(v_r(s) = e^r\). Now we need to express \(r, s\) in terms of \(x, y\). First, \(s = a^{-1}x\), and next \(r = y - bs = y - (b/a)x\). Thus, the solution of the PDE is

\[
u(x, y) = e^{y-(b/a)x} = e^y e^{-(b/a)x}.
\]

Note that the ODE along the characteristic in this case (vanishing \(c\)) stated that \(u\) is constant along characteristics. In general, the ODE along the characteristic propagates, or transports, the values of \(u\) on the initial curve, so this ODE may be called a transport equation.

Note also that we used \(a \neq 0\) in the solution. In general, while we can always solve the ODEs for small \(s\) at least, there is no guarantee that the map \((r, s) \mapsto (x, y)\) is invertible, or that the inverse map is differentiable. The problem is (locally) caused by integral curves \(\gamma_r(s)\) which at \(s = 0\) are tangent to \(\Gamma(r)\). For our problem, \(\Gamma\) was a parameterization of the \(y\)-axis, and thus we had to make sure that \(V\) is not tangent to the \(y\)-axis, i.e. has non-zero \(x\)-component, which is precisely the statement \(a \neq 0\). In general, we call the initial value problem \textit{non-characteristic} if \(V\) is not tangent to \(\Gamma\). Note that, now writing \(x = x(r, s), \ y = y(r, s),\)

\[
\frac{\partial x}{\partial r}(r, 0) = \Gamma_1'(r), \quad \frac{\partial y}{\partial r}(r, 0) = \Gamma_2'(r),
\]
since $x(r,0) = \gamma_1(r)$, $y(r,0) = \gamma_2(r)$, while from the equation of the characteristic curves,
\[
\frac{\partial x}{\partial s}(r,0) = a(\Gamma(r)), \quad \frac{\partial y}{\partial s}(r,0) = b(\Gamma(r)).
\]
The inverse function theorem tells us that the map $(r,s) \mapsto (x(r,s), y(r,s))$ is invertible near a point $(r_0,0)$ if the Jacobian matrix
\[
\begin{bmatrix}
\frac{\partial x}{\partial s} & \frac{\partial y}{\partial s}
\end{bmatrix}
\]
is invertible, i.e. if its rows are linearly independent, which is exactly the statement that $V$ is not tangent to the initial curve $\Gamma$.

Also, when $V$ is tangent to $\Gamma$, the difficulty is not an artifact of our method: the PDE tells us the derivative of $u$ along $V$, while the initial condition tells us the derivative of $u$ along $\Gamma$, so if $V$ is tangent to $\Gamma$, typically these two conditions will contradict each other.

A more interesting problem is the following:

**Problem 2.** Solve $u_x + yu_y = y^2$, with the initial condition $u(0,y) = \sin y$.

**Solution.** Now the initial curve is $\Gamma(r) = (0,r)$, the $y$-axis, so $\Gamma'(r) = (0,1)$, and $V = (1,y)$, so indeed this is a non-characteristic initial value problem. The equations for the characteristic curves are
\[
\begin{align*}
x'_r(s) &= 1, \quad x_r(0) = 0, \\
y'_r(s) &= y_r(s), \quad y_r(0) = r.
\end{align*}
\]
The solution is $x_r(s) = s$, $y_r(s) = re^s$. The ODE along the characteristic curves is
\[
v'_r(s) = y^2_r(s), \quad v_r(0) = \sin r.
\]
Rewriting $y$ in terms of $r, s$, and dropping the subscript $r$ to simplify notation,
\[
v'(s) = r^2e^{2s}, \quad v(0) = \sin r.
\]
The solution of the ODE is
\[
v(s) = \sin r + (r^2/2)(e^{2s} - 1)
\]
Now expressing $r, s$ in terms of $x, y$, $s = x$ while $r = ye^{-x} = ye^{-x}$. Thus,
\[
u(x, y) = \sin(ye^{-x}) + (1/2)y^2e^{-2x}(e^{2x} - 1) = \sin(ye^{-x}) + (1/2)y^2 - (1/2)y^2e^{-2x}.
\]

We can make this into a semilinear PDE by changing the right hand side:

**Problem 3.** Solve $u_x + yu_y = u^2$, with the initial condition $u(0,y) = \sin y$.

**Solution.** The characteristic curves are unchanged, as only the right hand side of the PDE was altered. Thus, $x_r(s) = s$, $y_r(s) = re^s$, and conversely $s = x$ while $r = ye^{-x}$. The ODE along the characteristic curve now is
\[
v'(s) = v(s)^2, \quad v(0) = \sin r.
\]
The solution is, from $\int_0^{s_0} \frac{v'}{v} ds = \int_0^{s_0} 1 ds = s_0$, $-\frac{1}{v_0} = s$, i.e. $\frac{1}{\sin r} - \frac{1}{v} = s$,
\[
v = \left(\frac{1}{2 + \sin r} - s\right)^{-1} = \frac{\sin r}{1 - s\sin r}.
\]
(Note that some of the algebraic manipulations above don’t work when $v = 0$, so one should check the result at the end of the calculations!) Thus,
\[
u(x, y) = \frac{\sin(ye^{-x})}{1 - xe^{-x}}.
\]
We remark that in this case the solution of the ODE along the characteristic curve blows up when the denominator vanishes (though it is non-zero for sufficiently small $x$, namely $|x| < 1$), which is fairly typical of non-linear ODE, and thus all one can expect is local solutions near the initial curve $\Gamma$ in general.

There is a slightly different way of looking at our characteristic curves. Namely, there is some arbitrariness in the way they are parameterized, since the PDE can be rewritten by multiplying through by any (non-zero) factor. In particular, if $a \neq 0$ near a point on $\Gamma$, we can divide by $a$, so the new PDE is

$$u_x + \frac{b(x, y)}{a(x, y)} u_y = \frac{c(x, y, u)}{a(x, y)}.$$  

(7)

In this version the vector field $V = (a, b)$ is replaced by $W = (1, b/a)$, which is in the same direction as $(a, b)$, so the integral curves have the same image, but are parameterized differently. The equations for the integral curves of $W$ are

$$x'(s) = 1, \quad y'(s) = b(x(s), y(s))/a(x(s), y(s)), \quad x(0) = \Gamma_1(r), \quad y(0) = \Gamma_2(r).$$

Thus, $x(s) = s + \Gamma_1(r)$, i.e. $x$ is simply a shifted version of the variable $s$. We might as well use $x$ as the variable along the integral curve instead, i.e. let $y = y(x)$, so

$$\frac{dy}{dx} = \frac{y'(s)}{x'(s)} = b(x, y)/a(x, y).$$

We can thus completely eliminate the variable $s$, and we only need to keep the parameter $r$, i.e. in full notation,

$$\frac{dy_r}{dx} = \frac{y'_r(s)}{x'_r(s)} = b(x, y_r(x))/a(x, y_r(x)).$$

If our initial curve is $\Gamma(r) = (\Gamma_1(r), \Gamma_2(r))$, then we need

$$y_r(\Gamma_1(r)) = \Gamma_2(r),$$

so that $\Gamma(r) = (\Gamma_1(r), y_r(\Gamma_1(r)))$. In addition, the initial condition for the ODE along the characteristic curves is

$$v_r(\Gamma_1(r)) = u(\Gamma_1(r), \Gamma_2(r)) = \phi(r).$$

Let us redo our Problem 2:

**Problem.** (Problem 2 restated.) Solve $u_x + y u_y = y^2$, with the initial condition $u(0, y) = \sin y$.

**Solution.** The initial curve is $\Gamma(r) = (0, r)$ as before, so the initial condition for characteristic curves is

$$y_r(0) = r.$$

Thus, the characteristic curves solve

$$\frac{dy_r}{dx} = y_r, \quad y_r(0) = r.$$

The solution is

$$y_r(x) = ye^{-x},$$

so $r = ye^{-x}$. The ODE along the characteristic curves now is

$$\frac{dv_r}{dx} = y_r^2, \quad v_r(0) = \sin r.$$

Substituting $y_r = re^x$, we get

$$\frac{dv_r}{dx} = r^2 e^{2x}, \quad v_r(0) = \sin r,$$
so
\[ u_r(x) = \sin r + (r^2/2)(e^{2x} - 1), \]
and thus
\[ u(x, y) = \sin(y e^{-x}) + (1/2)y^2 e^{-2x}(e^{2x} - 1) = \sin(y e^{-x}) + (1/2)y^2 - (1/2)y^2 e^{-2x}. \]

Of course, there are numerous variations one can do – for instance, one can ignore the initial conditions initially and instead find a ‘general solution’, and then try to match the initial conditions. This may be convenient if the initial condition is on a curve that is hard to handle – instead one can pretend that one has (unknown) initial conditions on a simple curve nearby, e.g. a coordinate axis, find the general solution in terms of these unknown initial data on the simple curve, and then match this general solution with the initial condition.

It is also important to realize that while here we discussed solving the scalar PDE in 2 independent variables, \( x \) and \( y \), the method works for \textit{any number} of independent variables \( x_1, \ldots, x_n \); of course the ODEs for the characteristic curves may be harder to solve then. Thus, writing \( x = (x_1, \ldots, x_n) \), for the PDE
\[ \sum_{j=1}^{n} a_j(x)u_{x_j}(x) = c(x, u(x)), \quad u(\Gamma(r)) = \phi(r), \]
where \( r = (r_1, \ldots, r_{n-1}) \) is an \( n-1 \)-dimensional variable parameterizing the initial \textit{hypersurface} \( \Gamma(r) = (\Gamma_1(r), \ldots, \Gamma_n(r)) \), the \( a_j \) are \( C^1 \) and real along with the \( \Gamma_j \), the vector field \( V \) becomes
\[ V = (a_1, \ldots, a_n), \]
and the equation for the characteristic curves is
\[ \frac{\partial x_j}{\partial s}(r, s) = a_j(x(r, s)), \quad j = 1, \ldots, n, \]
\[ x_j(r, 0) = \Gamma_j(r). \]
As above, \( v(r, s) = u(x(r, s)) \) satisfies
\[ \frac{\partial v}{\partial s}(r, s) = c(x(r, s), v(r, s)), \quad v(r, 0) = \phi(r). \]
Thus, the same methods as above work for the non-characteristic initial value problem, i.e. if \( V \) is not tangent to the hypersurface \( \Gamma \).

**Problem 4.** Solve \( u_x + yu_y + zu_z = u \), with the initial condition \( u(0, y, z) = y^2 + z^2 \).

\textit{Solution.} To avoid confusion, we’ll label the variables as \( (x_1, x_2, x_3) \). Now the initial surface is \( \Gamma(r_1, r_2) = (0, r_1, r_2) \), the \( x_2x_3 \)-plane, so the tangent plane of \( \Gamma \) is the span of \( (0, 1, 0) \) and \( (0, 0, 1) \), and \( V = (1, x_2, x_3) \), so indeed this is a non-characteristic initial value problem. The equations for the characteristic curves are
\[ x_1'(r, s) = 1, \quad x_1(r_1, r_2, 0) = 0, \]
\[ x_2'(r, s) = x_2(r, s), \quad x_2(r_1, r_2, 0) = r_1, \]
\[ x_3'(r, s) = x_3(r, s), \quad x_3(r_1, r_2, 0) = r_2. \]
The solution of the ODE is $v(t) = e^t$. The ODE along the characteristic curves is

$$v'(t) = v(t), \quad v(0) = e^0.$$ 

The solution of the ODE is $v(t) = e^t$. Now expressing $t$ in terms of $x$, $s = x$, while $r_1 = e^{-x}x_2 = e^{-x_1}x_2$, $r_2 = e^{-x}x_3 = e^{-x_1}x_3$. Thus,

$$u(x) = (e^{-2x_1}x_2^2 + e^{-2x_1}x_3^2)e^{-x_1} = (x_2^2 + x_3^2)e^{-x_1}.$$ 

Returning to the original notation,

$$u(x, y, z) = (y^2 + z^2)e^{-x}.$$ 

Although the title of this chapter is scalar equations, the theory is unchanged for systems of equations, provided the principal (i.e. first order) terms involve the same vector field. For instance, consider the system

$$V(x, y) \begin{pmatrix} u_1(x, y) \\ \vdots \\ u_N(x, y) \end{pmatrix} = \begin{pmatrix} F_1(x, y, u_1(x, y), \ldots, u_N(x, y)) \\ \vdots \\ F_N(x, y, u_1(x, y), \ldots, u_N(x, y)) \end{pmatrix},$$

with initial conditions

$$\begin{pmatrix} u_1(\Gamma(r)) \\ \vdots \\ u_N(\Gamma(r)) \end{pmatrix} = \begin{pmatrix} \phi_1(r) \\ \vdots \\ \phi_N(r) \end{pmatrix},$$

where the left hand side means of the PDE simply that we apply $V$ componentwise, i.e. we have

$$\begin{pmatrix} V(x, y)u_1(x, y) \\ \vdots \\ V(x, y)u_N(x, y) \end{pmatrix}$$

there. As before, we consider integral curves $\gamma(r, s) = \gamma_r(s) = (x(r, s), y(r, s))$ of $V$ through an initial curve $\Gamma = \Gamma(r)$; we find these solving the same ODEs as before, and the criterion for inverting the change of coordinates $(r, s) \mapsto (x(r, s), y(r, s))$ is still the non-tangency of $V(\Gamma(r))$ to $G$. Substituting in $v_j(r, s) = u_j(\gamma_r(s))$ yields the system

$$\begin{pmatrix} \frac{\partial}{\partial s} v_1(r, s) \\ \vdots \\ \frac{\partial}{\partial s} v_N(r, s) \end{pmatrix} = \begin{pmatrix} F_1(x(r, s), y(r, s), v_1(r, s), \ldots, v_N(r, s)) \\ \vdots \\ F_N(x(r, s), y(r, s), v_1(r, s), \ldots, v_N(r, s)) \end{pmatrix},$$

which is a system of ODEs (just like the system of ODEs for $(x, y)$ was already a system even in the scalar case!), and which is thus solvable at least for small $s$.

**Problem 5.** Solve the system of ODE $u_x + yu_y = w$, $w_x + yw_y = u$, with initial conditions $u(0, y) = y$, $w(0, y) = 1$.

**Solution.** The characteristic curves are the same as in Problem 2, so $x(r, s) = s$, $y(r, s) = re^s$, and the inverse map is $s = x$, $r = e^{-x}y$. The ODE for $(u, w)$, written as $(v_1, v_2)$, now is

$$\begin{pmatrix} \frac{\partial v_1}{\partial s} = v_2(r, s), & v_1(r, 0) = r \\ \frac{\partial v_2}{\partial s} = v_1(r, s), & v_2(r, 0) = 1 \end{pmatrix}.$$ 

Recall from ODE system theory that one should rewrite the right hand side as a linear system

$$\frac{\partial}{\partial s} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_2(r, 0) = \left( \begin{array}{c} 1 + s \\ 1 - s \end{array} \right) \end{pmatrix}.$$
The solution of this ODE system then is
\[
\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \exp \left( s \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) \begin{pmatrix} r \\ 1 \end{pmatrix}.
\]
Here the exponential can be calculated either by summing a Taylor series, namely it is
\[
\sum_{n=0}^{\infty} \frac{s^n}{n!} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} r \\ 1 \end{pmatrix},
\]
or by diagonalizing the matrix by conjugation by another matrix, and exponentiating the diagonal entries (i.e. the eigenvalues) of this matrix. In this case the eigenvectors are \( \begin{pmatrix} 1 \\ 1 \end{pmatrix} \) of eigenvalue 1 and \( \begin{pmatrix} 1 \\ -1 \end{pmatrix} \) with eigenvalue \(-1\), so the diagonalization is (using the eigenvectors as the columns in the first factor on the right hand side, the eigenvalues as the diagonal entries of the second factor)
\[
\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1},
\]
and the solution is thus (since \( \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \))
\[
\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} r \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^s & e^{-s} \\ e^s & -e^{-s} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} r \\ 1 \end{pmatrix}
\]
\[
= \begin{pmatrix} \cosh s & \sinh s \\ \sinh s & \cosh s \end{pmatrix} \begin{pmatrix} r \\ 1 \end{pmatrix} = \begin{pmatrix} r \cosh s + \sinh s \\ r \sinh s + \cosh s \end{pmatrix}.
\]
Substituting in \( r \) and \( s \) we finally obtain
\[
\begin{pmatrix} u(x, y) \\ w(x, y) \end{pmatrix} = \begin{pmatrix} ye^{-x} \cosh x + \sinh x \\ ye^{-x} \sinh x + \cosh x \end{pmatrix} = \begin{pmatrix} \frac{y}{2}(1 + e^{-2x}) + \sinh x \\ \frac{y}{2}(1 - e^{-2x}) + \cosh x \end{pmatrix}
\]
for the solution.

Here we assumed we were working in \( \mathbb{R}^2 \); changing to \( \mathbb{R}^n \) is purely a matter of notational complexity.