Before we start with the actual proof of convergence we need a small result about piece-wise linear interpolations. Remember, in the Semi Lagrangian scheme we have to evaluate $\tilde{v}(t,x)$, the velocity approximation, at x_D, which is not a grid-point. This requires to interpolate this value from values at grid points. The proof below requires the interpolation to fulfill the following:

$$|f(x) - g(y)| \leq \max_{x_i} |f(x_i) - g(x_i)| \tag{1}$$

where $f(x), g(y)$ are interpolations of the function f, g at x, y and the maximum is taken over all grid points. In order to realize such an interpolation one could use the following construction. Assume $\{x_0, x_1, \cdots, x_n\}$ are grid points which span the smallest convex set containing x. Then there are unique real numbers $t_i \geq 0$ with $\sum_{i=0}^n t_i = 1$ and $x = \sum_i t_i x_i$. We interpolate a function f, which values are known at grid points, by

$$f(x) := \sum_i t_i f(x_i)$$

If g is another function we compute

$$|f(x) - g(x)| = |\sum_i t_i f(x_i) - t_i g(x_i)| = |\sum_i t_i(f(x_i) - g(x_i))|$$

$$\leq |\sum_i t_i| \max_i |f(x_i) - g(x_i)| = \max_i |f(x_i) - g(x_i)|$$

Having that we come to the actual proof of convergence. We assume that F fulfills a Lipschitz-condition with constant $L > 0$ in all variables. Also, since F is continuous and our spatial region is bounded, we may assume F to be bounded by some $M > 0$. If $\tilde{v}$ is the solution obtained by iterating the scheme and v is the true solution, we show

$$\tilde{v}(t,x) \to v(t,x)$$

for all grid points x and t when $\Delta t \to 0$ and $\sup_x |\tilde{v}(0,x) - v(0,x)| \to 0$. The later expresses Δx to be chosen smaller and smaller. By the mean-value theorem applied on v we have for some $\tau \in [0, 1]$

$$v(t^{k+1}, x_A) = v(t^k, x_D) + \Delta t v'(t^k + \tau\Delta t, x_D)) = \\ v(t^k, x_D) + \Delta t F(t^k + \tau\Delta t, x_D, v(t^k, x_D)) \tag{2}$$

Note, x_D is the true departure point for the volume at x_A. By mean-value theorem applied on x it fulfills

$$x_D = x_A + \eta\Delta t x'(t^k + \eta\Delta t) \\ x_A + \eta\Delta t v(t^k + \eta\Delta t, x_D) \tag{3}$$

for some $\eta \in [-1, 0]$. Remember, the departure point from the iteration fulfills

$$\tilde{x}_D = \tilde{x}_A + \tilde{v}(t^k, \tilde{x}_D)$$

So we have, since $\tilde{x}_A = x_A$,

$$|x_D - \tilde{x}_D| = |\tilde{v}(t^k, \tilde{x}_D) - \eta\Delta t v(t^k + \eta\Delta t, x_D)| \tag{4} \\ \leq |\tilde{v}(t^k, \tilde{x}_D) - v(t^k, x_D)| + \sigma|\eta|\Delta t|v'(t_k + \sigma\eta\Delta t, x_D)| \\ = |\tilde{v}(t^k, \tilde{x}_D) - v(t^k, x_D)| + \sigma|\eta|\Delta t|F(t_k + \sigma\eta\Delta t, x_D, v(t_k, x_D))|$$

where again we applied the mean-value theorem with suitable $\sigma \in [0, 1]$ such that

$$v(t^k + \eta\Delta t) = v(t^k) + \eta\Delta t v'(t^k + \sigma\eta\Delta t)$$

Next we subtract $v(t^{k+1}, x_A)$ from $\tilde{v}(t^{k+1}, x_A)$, which gives

$$\tilde{v}(t^{k+1}, x_A) - v(t^{k+1}, x_A) = \tilde{v}(t^k, \tilde{x}_D) - v(t^k, x_D) + \\ \Delta t F(t^k, \tilde{x}_D, \tilde{v}(t^k, \tilde{x}_D)) - \Delta t F(t^k + \tau\Delta t, x_D, v(t^k, x_D))$$

We can estimate this by using the triangle inequality and 'extending by zeros':

$$|\tilde{v}(t^{k+1}, x_A) - v(t^{k+1}, x_A)| \leq |\tilde{v}(t^k, \tilde{x}_D) - v(t^k, x_D)| + \\ |\Delta t F(t^k, \tilde{x}_D, \tilde{v}(t^k, \tilde{x}_D)) - \Delta t F(t^k + \tau\Delta t, \tilde{x}_D, \tilde{v}(t^k, \tilde{x}_D))| \\ + |\Delta t F(t^k + \tau\Delta t, \tilde{x}_D, \tilde{v}(t^k, \tilde{x}_D)) - \Delta t F(t^k + \tau\Delta t, x_D, \tilde{v}(t^k, \tilde{x}_D))| \\ + |\Delta t F(t^k + \tau\Delta t, x_D, \tilde{v}(t^k, \tilde{x}_D)) - \Delta t F(t^k + \tau\Delta t, x_D, v(t^k, x_D))|$$

Using the Lipschitz-condition of F yields

$$|\tilde{v}(t^{k+1}, x_A) - v(t^{k+1}, x_A)| \leq |\tilde{v}(t^k, x_D) - v(t^k, x_D)| + (\Delta t)^2 \tau L$$
$$+ \Delta t \tau L |\tilde{x}_D - x_D| + \Delta t L |\tilde{v}(t^k, \tilde{x}_D) - v(t^k, x_D)|$$

Using (4) gives

$$|\tilde{v}(t^{k+1}, x_A) - v(t^{k+1}, x_A)| \leq |\tilde{v}(t^k, x_D) - v(t^k, x_D)| + (\Delta t)^2 \tau L$$
$$+ \Delta t \tau L |\tilde{v}(t^k, \tilde{x}_D) - v(t^k, x_D)|$$
$$+ \tau \sigma |\eta| (\Delta t)^2 L |F(t_k + \sigma \eta \Delta t, x_D, v(t_k, x_D))| +$$
$$\Delta t |\tilde{v}(t^k, \tilde{x}_D) - v(t^k, x_D)|$$

Next we define for any n

$$e_n := max_x |\tilde{v}(t^n, x) - v(t^n, x)|$$

where the maximum is taken over all grid-points. Using this and the bound of F in above estimate gives

$$|\tilde{v}(t^{k+1}, x_A) - v(t^{k+1}, x_A)| \leq e_k + (\Delta t)^2 \tau L + \Delta t \tau L$$
$$e_k + \tau \sigma |\eta| (\Delta t)^2 L M + \Delta t e_k$$

Note, how we have used (1) since x_D is not a grid point. Since x_A was chosen arbitrarily and the r.h.s does not depend on any grid-point, we obtain by taking the maximum over all grid-points

$$e_{k+1} \leq e_k + (\Delta t)^2 \tau L + \Delta t \tau L e_k + \tau \sigma |\eta| (\Delta t)^2 L M + \Delta t e_k$$

The remaining part is a very standard argument used in convergence proofs for ODE's. We arrived at an estimate of the form

$$e_{k+1} \leq (1 + \alpha \Delta t) e_k + \beta (\Delta t)^2$$

The r.h.s can be transformed into

$$(1 + \alpha \Delta t) \left(\frac{\beta \Delta t}{\alpha} + e_k \right) - \frac{\beta \Delta t}{\alpha}$$

So we obtain

$$e_{k+1} + \frac{\beta \Delta t}{\alpha} \leq (1 + \alpha \Delta t) \left(\frac{\beta \Delta t}{\alpha} + e_k \right)$$

and by estimating $1 + \alpha \Delta t \leq exp(\alpha \Delta t)$

$$e_{k+1} + \frac{\beta \Delta t}{\alpha} \leq exp(\alpha \Delta t) \left(\frac{\beta \Delta t}{\alpha} + e_k \right)$$

We further can apply this later estimate recursively on its r.h.s to get

$$e_{k+1} + \frac{\beta \Delta t}{\alpha} \leq exp((k+1)\alpha \Delta t) \left(\frac{\beta \Delta t}{\alpha} + e_0 \right)$$

By using $(k+1)\Delta t \leq T$, this finally yields

$$e_{k+1} \leq exp(T\alpha) \left(\frac{\beta \Delta t}{\alpha} + e_0 \right) - \frac{\beta \Delta t}{\alpha}$$

This shows, for any $n > 0$, $e_n \to 0$ when $\Delta t \to 0$ and $e_0 \to 0$. Note, the convergence of e_0 to 0, by definition of e_0, depends on the spacial resolution Δx. Altogether this shows, $\tilde{v}$ approaches the true v when $\Delta t, \Delta x$ approach 0.