A DIFFERENCE SCHEME FOR SOLVING AN INITIAL VALUE PROBLEM FROM SEMICONDUCTOR DEVICE THEORY*1)

KUO PEN-YU (郭本瑜)

(*Shanghai University of Science and Technology, Shanghai, China*)

I. Introduction

We consider the following problem from the semiconductor device theory:

$$
\begin{aligned}
&\frac{\partial U}{\partial t}=\nu\sum_{i=1}^{n}\frac{\partial^2 U}{\partial x_i^2}-\nabla\cdot(U\nabla\psi)-R(U,V), \quad (x,t)\in\Omega\times(0,T],\\
&\frac{\partial V}{\partial t}=\nu\sum_{i=1}^{n}\frac{\partial^2 V}{\partial x_i^2}+\nabla\cdot(V\nabla\psi)-R(U,V), \quad (x,t)\in\Omega\times(0,T],\\
&p\sum_{i=1}^{n}\frac{\partial^2 \psi}{\partial x_i^2}=U-V-N, \quad (x,t)\in\Omega\times(0,T],\\
&\frac{\partial U}{\partial n}=\frac{\partial V}{\partial n}=\frac{\partial \psi}{\partial n}=0, \quad (x,t)\in\Gamma\times[0,T],\\
&U(x,0)=U_0(x), \quad x\in\Omega,\\
&V(x,0)=V_0(x), \quad x\in\Omega,
\end{aligned}
\tag{1.1}
$$

where $x=(x_1, x_2, \cdots, x_n)$, $\Omega=\{X\,|\,0<x_j<L,\ 1\leqslant j\leqslant n\}$, Γ is the boundary of Ω, T is a specified positive constant, ν, p, q are positive constants, N is a specified Hölder continuous function of x, t,

$$R(U,V)=\frac{UV-1}{q(U+V+2)}, \tag{1.2}$$

$U_0(x), V_0(x)$ are twice continuously differentiable in x and strictly positive in $\Omega+\Gamma$, and

$$\int_\Omega (U_0(x)-V_0(x)-N(x))dx=0. \tag{1.3}$$

By a solution, we mean a set of three functions U, V, ψ of (x, t) in $\Omega\times[0, T]$, twice continuously differentiable in x and continuously differentiable in t satisfying (1.1)—(1.3), with U and V positive. For uniqueness we require also

$$\int_\Omega \psi(x,t)dx=0, \quad \forall t\in[0,T].$$

Mock[1] proved that, under the above conditions, (1.1) has a unique solution, and gave a difference scheme for solving (1.1), but without the proof of convergence.

In this paper we give a scheme for solving (1,1) with a strict proof of its convergence.

* Received September 10, 1982.

1) This work is suggested by Professor R. Glowinski.

II. Notations and Lemmas

Let h be the mesh size in variable x_j, $j=1, 2, \cdots, n$. Q denotes a mesh point, and e_j is a unit vector, i. e.

$$e_j=(\underbrace{0, 0, \cdots, 0}_{j-1}, 1, \underbrace{0, \cdots, 0}_{n-j})^T.$$

Ω_h denotes the set of internal mesh points. Γ_h is the boundary of Ω_h,

$$\Gamma_{jM}=\{Q|Q\in\Gamma_h,\ Q-he_j\in\Omega_h\},$$
$$\Gamma_{jm}=\{Q|Q\in\Gamma_h,\ Q+he_j\in\Omega_h\},$$
$$\Gamma_j=\Gamma_{jM}+\Gamma_{jm}.$$

τ denotes the mesh size of variable t, $\lambda=\tau h^{-2}$.

Let η be the discrete function. $\eta(Q, k)$ denotes the value of η at point Q and time $t=k\tau$.

$$\eta(k)=\{\eta(Q, k)/Q\in\Omega_h+\Gamma_h\}.$$

For simplicity, we denote $\eta(Q, k)$ by $\eta(Q)$ or $\eta(k)$. We define

$$\eta_{x_j}(Q, k)=\frac{1}{h}[\eta(Q+he_j, k)-\eta(Q, k)],$$

$$\eta_{\bar{x}_j}(Q, k)=\frac{1}{h}[\eta(Q, k)-\eta(Q-he_j, k)],$$

$$\eta_n(Q, k)=\begin{cases}\eta_{\bar{x}_j}(Q, k), & \text{if } Q\in\Gamma_{jM},\\ -\eta_{x_j}(Q, k), & \text{if } Q\in\Gamma_{jm},\end{cases}$$

$$\Delta_{x_j}\eta(Q, k)=\eta_{\bar{x}_jx_j}(Q, k), \quad \Delta\eta(Q, k)=\sum_{j=1}^{n}\Delta x_j\eta(Q, k),$$

$$\eta_t(Q, k)=\frac{1}{\tau}[\eta(Q, k+1)-\eta(Q, k)],$$

and define the following scalar product and norms

$$(\eta, \xi)=\sum_{Q\in\Omega_h}h^n\eta(Q)\xi(Q),$$

$$\|\eta\|^2=(\eta, \eta),$$

$$\|\eta\|_1^2=\frac{1}{2}\sum_{j=1}^{n}(\|\eta_{x_j}\|^2+\|\eta_{\bar{x}_j}\|^2),$$

$$|\eta|_{\Gamma_h}=\max_{Q\in\Gamma_h}|\eta(Q)|,$$

$$\|\eta\|_{\Gamma_h}^2=\sum_{Q\in\Gamma_h}h^{n-1}\eta^2(Q).$$

We will use the following lemmas.

Lemma 1. $2(\eta, \eta_t)=(\|\eta\|^2)_t-\tau\|\eta_t\|^2$.

Lemma 2.

$$(\eta, \xi_{x_j})+(\xi, \eta_{\bar{x}_j})=h^{n-1}\sum_{Q\in\Gamma_{jM}}\eta(Q-he_j)\xi(Q)-h^{n-1}\sum_{Q\in\Gamma_{jm}}\eta(Q)\xi(Q+he_j), \quad (2.1)$$

$$(\eta, \xi_{\bar{x}_j})+(\xi, \eta_{x_j})=h^{n-1}\sum_{Q\in\Gamma_{jM}}\eta(Q)\xi(Q-he_j)-h^{n-1}\sum_{Q\in\Gamma_{jm}}\eta(Q+he_j)\xi(Q). \quad (2.2)$$

The proofs come from Abel's formula directly.

Lemma 3.

$$2(\eta,\ \Delta\xi)+\sum_{j=1}^{n}[(\eta_{x_j},\ \xi_{x_j})+(\eta_{\bar{x}_j},\ \xi_{\bar{x}_j})]$$
$$=2h^{n-1}\sum_{Q\in\Gamma_h}\eta(Q)\xi_n(Q)-h^n\sum_{Q\in\Gamma_h}\eta_n(Q)\xi_n(Q).$$

Proof. From (2.1), we have

$$(\eta,\ \Delta_{x_j}\xi)+(\xi_{\bar{x}_j},\ \eta_{\bar{x}_j})=h^{n-1}\sum_{Q\in\Gamma_{jM}}\eta(Q-he_j)\xi_{\bar{x}_j}(Q)-h^{n-1}\sum_{Q\in\Gamma_{jm}}\eta(Q)\xi_{\bar{x}_j}(Q+he_j). \tag{2.3}$$

For $Q\in\Gamma_{jm}$, we have

$$\xi_{\bar{x}_j}(Q+e_jh)=-\xi_n(Q).$$

Hence

$$(\eta,\ \Delta_{x_j}\xi)+(\xi_{\bar{x}_j},\ \eta_{\bar{x}_j})=h^{n-1}\sum_{Q\in\Gamma_{jM}}\eta(Q-he_j)\xi_n(Q)+h^{n-1}\sum_{Q\in\Gamma_{jm}}\eta(Q)\xi_n(Q). \tag{2.4}$$

From (2.2), we have

$$(\eta,\ \Delta_{x_j}\xi)+(\xi_{x_j},\ \eta_{x_j})=h^{n-1}\sum_{Q\in\Gamma_{jM}}\eta(Q)\xi_{x_j}(Q-he_j)-h^{n-1}\sum_{Q\in\Gamma_{jm}}\eta(Q+he_j)\xi_{x_j}(Q) \tag{2.5}$$

and

$$(\eta,\ \Delta_{x_j}\xi)+(\xi_{x_j},\ \eta_{x_j})=h^{n-1}\sum_{Q\in\Gamma_{jM}}\eta(Q)\xi_n(Q)+h^{n-1}\sum_{Q\in\Gamma_{jm}}\eta(Q+he_j)\xi_n(Q). \tag{2.6}$$

By summing (2.4) and (2.6) for $j=1,\ 2,\ \cdots,\ n$, we have

$$2(\eta,\ \Delta\xi)+\sum_{j=1}^{n}[(\eta_{x_j},\ \xi_{x_j})+(\eta_{\bar{x}_j},\ \xi_{\bar{x}_j})]$$
$$=h^{n-1}\sum_{j=1}^{n}\{\sum_{Q\in\Gamma_{jM}}[\eta(Q)+\eta(Q-he_j)]\xi_n(Q)$$
$$+\sum_{Q\in\Gamma_{jm}}[\eta(Q)+\eta(Q+he_j)]\xi_n(Q)\}. \tag{2.7}$$

Since $\quad \eta(Q-he_j)\xi_n(Q)=-h\eta_n(Q)\xi_n(Q)+\eta(Q)\xi_n(Q),\quad$ for $Q\in\Gamma_{jM}$

and $\quad \eta(Q+he_j)\xi_n(Q)=-h\eta_n(Q)\xi_n(Q)+\eta(Q)\xi_n(Q),\quad$ for $Q\in\Gamma_{jm}$.

(2.7) gives the conclusion.

Especially, if $\eta_n=g$, then

$$2(\eta,\ \Delta\eta)+2\|\eta\|_1^2=2h^{n-1}\sum_{Q\in\Gamma_h}\eta(Q)g(Q)-h^n\sum_{Q\in\Gamma_h}g^2(Q), \tag{2.8}$$

$$2(\eta_t,\Delta\eta_t)+2\|\eta_t\|_1^2=2h^{n-1}\sum_{Q\in\Gamma_h}\eta_t g_t(Q)-h^n\sum_{Q\in\Gamma_h}g_t^2(Q). \tag{2.9}$$

Lemma 4. *If $\eta_n=g$, then*

$$2(\eta,\ \Delta\eta_t)+(\|\eta\|_1^2)_t-\tau\|\eta_t\|_1^2=2h^{n-1}\sum_{Q\in\Gamma_h}\eta(Q)g_t(Q)-h^n\sum_{Q\in\Gamma_h}g(Q)g_t(Q).$$

Proof. By putting $\xi=\eta_t$ in Lemma 3, we obtain

$$2(\eta,\ \Delta\eta_t)+\sum_{j=1}^{n}[(\eta_{x_j},\ \eta_{tx_j})+(\eta_{\bar{x}_j},\ \eta_{t\bar{x}_j})]=2h^{n-1}\sum_{Q\in\Gamma_h}\eta(Q)g_t(Q)-h^n\sum_{Q\in\Gamma_h}g(Q)g_t(Q).$$

From Lemma 1, we have

$$\sum_{j=1}^{n}[(\eta_{x_j},\ \eta_{tx_j})+(\eta_{\bar{x}_j},\ \eta_{t\bar{x}_j})]=(\|\eta\|_1^2)_t-\tau\|\eta_t\|_1^2.$$

Lemma 5. *If $\eta_n=g$, then*

$$2(\eta_t,\ \Delta\eta)+(\|\eta\|_1^2)_t-\tau\|\eta_t\|_1^2=2h^{n-1}\sum_{Q\in\Gamma_h}\eta_t(Q)g(Q)-h^n\sum_{Q\in\Gamma_h}g(Q)g_t(Q).$$

Lemma 6[2]**.** $\|\eta\xi\|^2\leq h^{-n}\|\eta\|^2\|\xi\|^2$.

Lemma 7. *If Γ_h is suitably regular, $s_1>0$ and h is sufficiently small, then*

$$\|\eta\|_{\Gamma_h}^2\leqslant\varepsilon_1\|\eta\|_1^2+C_1\left(1+\frac{1}{\varepsilon_1}\right)\|\eta\|^2,$$

where C_1 is a constant independent of η.

Proof. For simplicity, we suppose $n=2$, and $Mh=L-h$, M being an integer. Let Q denote a mesh point on Γ_{2m} with coordinates

$$x_1(Q)=ih,\quad x_2(Q)=0.$$

$\eta(ih,\ lh)$ denotes the value of η at point $x_1=ih$, $x_2=lh$. Then

$$\begin{aligned}\eta^2(ih,\ 0)&=\eta^2(ih,\ lh)-h\sum_{m=0}^{l-1}(\eta(ih,\ mh)+\eta(ih,\ mh+h))\eta_{x_2}(ih,\ mh)\\&\leqslant\left(1+\frac{h}{2\varepsilon}\right)\eta^2(ih,\ lh)+\frac{h}{2\varepsilon}\eta^2(ih,\ 0)+\varepsilon h\sum_{m=0}^{l-1}\eta_{x_2}^2(ih,\ mh)\\&\quad+\frac{h}{\varepsilon}\sum_{m=1}^{l-1}\eta^2(ih,\ mh).\end{aligned}$$

Hence

$$\left(1-\frac{h}{2\varepsilon}\right)\eta^2(ih,\ 0)\leqslant\left(1+\frac{h}{2\varepsilon}\right)\eta^2(ih,\ lh)+\varepsilon h\sum_{m=0}^{M}\eta_{x_2}^2(ih,\ mh)+\frac{h}{\varepsilon}\sum_{m=1}^{M}\eta^2(ih,\ mh).$$

By summing the above formula for all l, we have

$$L\left(1-\frac{h}{2\varepsilon}\right)\eta^2(ih,\ 0)\leqslant h\left(1+\frac{h+2L}{2\varepsilon}\right)\sum_{m=1}^{M}\eta^2(ih,\ mh)+\varepsilon Lh\sum_{m=0}^{M}\eta_{x_2}^2(ih,\ mh).$$

By summing the above formula for all points on Γ_{2m}, we get

$$L\left(1-\frac{h}{2\varepsilon}\right)\sum_{Q\in\Gamma_{2m}}h\eta^2(Q)\leqslant\left(1+\frac{h+2L}{2\varepsilon}\right)\|\eta\|^2+\varepsilon L\|\eta_{x_2}\|^2.$$

The rest of the proof is clear.

Remark 1. We also have

$$h^{n-1}\sum_{Q\in\Gamma_{jM}}\eta^2(Q-he_j)+h^{n-1}\sum_{Q\in\Gamma_{jm}}\eta^2(Q+he_j)\leqslant\varepsilon_1\|\eta\|_1^2+C_1\left(1+\frac{1}{\varepsilon_1}\right)\|\eta\|^2.$$

Lemma 8. *If $(1,\ \eta)=0$, then*

$$\|\eta\|^2\leqslant C_2\|\eta\|_1^2,$$

where C_2 is a positive constant depending on L.

Lemma 9. *If* (1) $(1,\ \eta)=0$,

(2) *L is so small that* $\sqrt{C_1C_2}<\sqrt{\frac{5}{2}}-1$,

(3) $\eta_n=g$, *for* $Q\in\Gamma_h$,

then

$$\|\eta\|_1^2\leqslant C_3(\|g\|_{\Gamma_h}^2+\|\Delta\eta\|^2),$$

where C_3 is a positive constant independent of η.

Proof. From (2.8) and Lemmas 7, 8, we get

$$\begin{aligned}2\|\eta\|_1^2&=-2(\eta,\ \Delta\eta)+2h^{n-1}\sum_{Q\in\Gamma_h}\eta(Q)g(Q)-h^n\sum_{Q\in\Gamma_h}g^2(Q)\\&\leqslant\frac{\|\eta\|^2}{2C_2}+2C_2\|\Delta\eta\|^2+\|\eta\|_{\Gamma_h}^2+2\|g\|_{\Gamma_h}^2\\&\leqslant\frac{1}{2}\|\eta\|_1^2+2C_2\|\Delta\eta\|^2+\varepsilon_1\|\eta\|_1^2+C_1\left(1+\frac{1}{\varepsilon_1}\right)\|\eta\|^2+2\|g\|_{\Gamma_h}^2\\&\leqslant\left(\frac{1}{2}+\varepsilon_1+C_1C_2\left(1+\frac{1}{\varepsilon_1}\right)\right)\|\eta\|_1^2+2C_2\|\Delta\eta\|^2+2\|g\|_{\Gamma_h}^2.\end{aligned}$$

By putting $\varepsilon_1=\sqrt{C_1C_2}$, we get the conclusion.

Lemma 10[8]. *If the following conditions hold:*

(1) $\eta(k)$ *is a nonnegative mesh function*, τ, M_0, M_1 *and* ρ *are positive constants,*

(2) $M(z)$, $f^*(z)$ *are such mesh functions that if* $z\in M_1h^m$, *then*

$$f^*(z)\leqslant 0,\quad M(z)\leqslant M_0 z,$$

(3) $\eta(k)\leqslant\rho+\tau\sum_{j=0}^{k-1}[M(\eta(j))+f^*(\eta(j))]$, *for* $k\geqslant 1$,

(4) $\eta(0)\leqslant\rho$,

(5) $\rho e^{t_0M_0}\leqslant M_1h^m$, $k\tau<t_0$,

then

$$\eta(k)\leqslant\rho e^{M_0k\tau}$$

III. Difference Scheme

We rewrite (1.1) as

$$\begin{cases}\dfrac{\partial U}{\partial t}=\nu\sum_{i=1}^{n}\dfrac{\partial^2U}{\partial x_i^2}-\dfrac{1}{2}\nabla\cdot(U\nabla\psi)-\dfrac{1}{2}(\nabla U\cdot\nabla\psi)+R_1(U,V)-R(U,V),\\ \dfrac{\partial V}{\partial t}=\nu\sum_{i=1}^{n}\dfrac{\partial^2V}{\partial x_i^2}+\dfrac{1}{2}\nabla\cdot(V\nabla\psi)+\dfrac{1}{2}(\nabla V\cdot\nabla\psi)+R_2(U,V)-R(U,V),\\ p\sum_{i=1}^{n}\dfrac{\partial^2\psi}{\partial x_i^2}=U-V-N,\end{cases}\tag{3.1}$$

where $R_1(U,V)=\dfrac{-U}{2p}(U-V-N)$, $R_2(U,V)=\dfrac{V}{2p}(U-V-N)$. In order to approximate (3.1), we define the difference operators

$$J(\eta,\varphi)=\frac{1}{4}\sum_{i=1}^{n}((\eta\varphi_{\bar{x}_i})_{x_i}+(\eta\varphi_{x_i})_{\bar{x}_i}+\eta_{\bar{x}_i}\psi_{\bar{x}_i}+\eta_{x_i}\varphi_{x_i})$$

and

$$A(\eta,\xi,\varphi_n)=\frac{h^{n-1}}{2}\sum_{i=1}^{n}\sum_{Q\in\Gamma_{iM}}(\eta(Q)\xi(Q-he_i)+\xi(Q)\eta(Q-he_i))\varphi_n(Q)$$
$$+\frac{h^{n-1}}{2}\sum_{i=1}^{n}\sum_{Q\in\Gamma_{im}}(\eta(Q)\xi(Q+he_i)+\xi(Q)\eta(Q+he_i))\varphi_n(Q).$$

We have

$$(J(\eta,\varphi),\xi)+(J(\xi,\varphi),\eta)=A(\eta,\xi,\varphi_n).\tag{3.2}$$

Proof. From (2.1) and (2.2), we have

$$((\eta\varphi_{\bar{x}_i})_{x_i},\xi)+(\eta\varphi_{\bar{x}_i},\xi_{\bar{x}_i})$$
$$=h^{n-1}\sum_{Q\in\Gamma_{iM}}\eta(Q)\varphi_{\bar{x}_i}(Q)\xi(Q-he_i)-h^{n-1}\sum_{Q\in\Gamma_{im}}\xi(Q)\varphi_{\bar{x}_i}(Q+he_i)\eta(Q+he_i),$$

and

$$((\eta\varphi_{x_i})_{\bar{x}_i},\xi)+(\eta\varphi_{x_i},\xi_{x_i})$$
$$=h^{n-1}\sum_{Q\in\Gamma_{iM}}\xi(Q)\eta(Q-he_i)\varphi_{x_i}(Q-he_i)-h^{n-1}\sum_{Q\in\Gamma_{im}}\xi(Q+he_i)\eta(Q)\varphi_{x_i}(Q).$$

Similarly we get

$$((\xi\varphi_{\bar{x}_i})_{x_i},\eta)+(\xi\varphi_{\bar{x}_i},\eta_{\bar{x}_i})$$
$$=h^{n-1}\sum_{Q\in\Gamma_{iM}}\xi(Q)\varphi_{\bar{x}_i}(Q)\eta(Q-he_i)-h^{n-1}\sum_{Q\in\Gamma_{im}}\eta(Q)\varphi_{\bar{x}_i}(Q+he_i)\xi(Q+he_i)$$

and

$$((\xi\varphi_{x_i})_{\bar{x}_i},\eta)+(\xi\varphi_{x_i},\eta_{x_i})$$
$$=h^{n-1}\sum_{Q\in\Gamma_{iM}}\eta(Q)\xi(Q-he_i)\varphi_{x_i}(Q-he_i)-h^{n-1}\sum_{Q\in\Gamma_{im}}\eta(Q+he_i)\xi(Q)\varphi_{x_i}(Q).$$

By summing the above four formulas and noticing the definition of φ_n, we get (3.2).

If $\varphi_n=g$ then from (3.2) we have

$$(J(\eta,\ \phi),\ \eta)=\frac{1}{2}A(\eta,\ \eta,\ g). \tag{3.3}$$

Especially, if $g=0$, we obatain

$$(J(\eta,\ \varphi),\ \eta)=0. \tag{3.4}$$

Now we denote the approximate solution of U, V and ψ, by u, v and φ. We construct the following scheme for solving (3.1)

$$\begin{cases} u_t(Q,\ k)-\nu\Delta(u+\sigma\tau u_t)(Q,\ k)+J(u+\delta\tau u_t,\ \varphi)(Q,\ k) \\ \qquad -R_1(u,\ v)(Q,\ k)+R(u,\ v)(Q,\ k)=0, \\ v_t(Q,\ k)-\nu\Delta(v+\sigma\tau v_t)(Q,\ k)-J(v+\delta\tau v_t,\ \varphi)(Q,\ k) \\ \qquad -R_2(u,\ v)(Q,\ k)+R(u,\ v)(Q,\ k)=0, \\ p\Delta\varphi(Q,\ k)-u(Q,\ k)+v(Q,\ k)+N(Q,\ k)=0, \\ (1,\ \varphi(k))=0, \\ u_n(Q,\ k)=0, \quad \text{for } Q\in\Gamma_h, \\ v_n(Q,\ k)=0, \quad \text{for } Q\in\Gamma_h, \\ \varphi_n(Q,\ k)=0, \quad \text{for } Q\in\Gamma_h, \\ u(Q,\ 0)=u_0(Q),\ v(Q,\ 0)=v_0(Q), \end{cases} \tag{3.5}$$

where $0\leqslant\delta\leqslant1$, $0\leqslant\sigma\leqslant1$. If $\sigma=\delta=0$, then (3.5) is an explicit scheme. Otherwise it is an implicit scheme.

IV. Error Estimation

Let $u=U+\tilde{u}$, $v=V+\tilde{v}$, $\varphi=\psi+\tilde{\varphi}$, and $\tilde{f}_i$ and $\tilde{g}_i$ be truncation errors. Then we obtain the following error equations

$$\begin{cases} \tilde{u}_t=\nu\Delta(\tilde{u}+\sigma\tau\tilde{u}_t)-J(\tilde{u},\ \psi+\tilde{\varphi})-\delta\tau J(\tilde{u}_t,\ \psi+\tilde{\varphi})-J(U+\delta\tau U_t,\ \tilde{\varphi}) \\ \qquad +R_1(U+\tilde{u},\ V+\tilde{v})-R_1(U,\ V)-R(U+\tilde{u},\ V+\tilde{v})+R(U,\ V)-\tilde{f}_1, \\ \tilde{v}_t=\nu\Delta(\tilde{v}+\sigma\tau\tilde{v}_t)+J(\tilde{v},\ \psi+\tilde{\varphi})+\delta\tau J(\tilde{v}_t,\ \psi+\tilde{\varphi})+J(V+\delta\tau V_t,\ \tilde{\varphi}) \\ \qquad +R_2(U+\tilde{u},\ V+\tilde{v})-R_2(U,\ V)-R(U+\tilde{u},\ V+\tilde{v})+R(U,\ V)-\tilde{f}_2, \\ p\Delta\tilde{\varphi}=\tilde{u}-\tilde{v}-\tilde{f}_3, \\ (1,\ \tilde{\varphi})=0, \\ \tilde{u}_n=\tilde{g}_1,\ \tilde{v}_n=\tilde{g}_2,\ \tilde{\varphi}_n=\tilde{g}_3, \quad \text{for } Q\in\Gamma_h, \\ \tilde{u}(0)=\tilde{u}_0,\ \tilde{v}(0)=\tilde{v}_0. \end{cases} \tag{4.1}$$

Taking the scalar product of the first formula of (4.1) with $2\tilde{u}$, from Lemmas 1—4 and (3.2), (3.3), we have

$$\begin{aligned} &\|\tilde{u}\|_t^2-\tau\|\tilde{u}_t\|^2+2\nu\|\tilde{u}\|_1^2+\nu\sigma\tau(\|\tilde{u}\|_1^2)_t-\nu\sigma\tau^2\|\tilde{u}_t\|_1^2 \\ &\quad =2\nu h^{n-1}\sum_{Q\in\Gamma_h}\tilde{u}(Q)\tilde{g}_1(Q)-\nu h^n\sum_{Q\in\Gamma_h}\tilde{g}_1^2(Q)+2\sigma\nu\tau h^{n-1}\sum_{Q\in\Gamma_h}\tilde{u}(Q)\tilde{g}_{1t}(Q) \\ &\qquad -\sigma\nu\tau h^n\sum_{Q\in\Gamma_h}\tilde{g}_1(Q)\tilde{g}_{1t}(Q)-A(\tilde{u},\ \tilde{u},\ \psi_n+\tilde{\varphi}_n)+2\delta\tau(\tilde{u}_t,\ J(\tilde{u},\ \psi+\tilde{\varphi})) \\ &\qquad -2\delta\tau A(\tilde{u}_t,\ \tilde{u},\ \psi_n+\tilde{\varphi}_n)-(2\tilde{u},\ J(v+\delta\tau U_t,\ \tilde{\varphi})-R_1(U+\tilde{u},\ V+\tilde{v}) \\ &\qquad +R_1(U,\ V)+R(U+\tilde{u},\ V+\tilde{v})-R(U,\ V)+\tilde{f}_1). \end{aligned} \tag{4.2}$$

Let m be a positive constant to be chosen later. By taking the scalar product of the first formula of (4.1) with $m\tau\tilde{u}_t$, from Lemmas 3—5 and (3.2), we have

$$m\tau\|\tilde{u}_t\|^2+\frac{m\nu\tau}{2}(\|\tilde{u}\|_1^2)_t-\frac{\nu m\tau^2}{2}\|\tilde{u}_t\|_1^2+\sigma\nu m\tau^2\|\tilde{u}_t\|_1^2$$

$$=\nu m\tau h^{n-1}\sum_{Q\in\Gamma_h}\tilde{u}_t(Q)\tilde{g}_1(Q)-\frac{\nu m\tau h^n}{2}\sum_{Q\in\Gamma_h}\tilde{g}_{1t}(Q)\tilde{g}_1(Q)$$

$$+\nu m\sigma\tau^2h^{n-1}\sum_{Q\in\Gamma_h}\tilde{u}_t(Q)\tilde{g}_{1t}(Q)-\frac{\nu m\sigma\tau^2h^n}{2}\sum_{Q\in\Gamma_h}\tilde{g}_{1t}^2(Q)$$

$$-\frac{\delta m\tau^2}{2}A(\tilde{u}_t,\ \tilde{u}_t,\ \psi_n+\tilde{\varphi}_n)-m\tau(\tilde{u}_t,\ J(\tilde{u},\ \psi+\tilde{\varphi}))$$

$$-(m\tau\tilde{u}_t,\ J(U+\delta\tau U_t,\ \tilde{\varphi})-R_1(U+\tilde{u},\ V+\tilde{v})+R_1(U,\ V)$$

$$+R(U+\tilde{u},\ V+\tilde{v})-R(U,\ V)+\tilde{f}_1). \tag{4.3}$$

By combining (4.2) with (4.3), we get

$$\|\tilde{u}\|_t^2+\tau(m-1)\|\tilde{u}_t\|^2+2\nu\|\tilde{u}\|_1^2+\nu\tau\left(\sigma+\frac{m}{2}\right)(\|\tilde{u}\|_1^2)_t+\nu\tau^2\left(\sigma m-\frac{m}{2}-\sigma\right)\|\tilde{u}_t\|_1^2$$

$$=\tau(2\delta-m)(\tilde{u}_t,\ J(\tilde{u},\ \psi+\tilde{\varphi}))-A(\tilde{u},\ \tilde{u},\ \psi_n+\tilde{\varphi}_n)-\frac{\delta m\tau^2}{2}A(\tilde{u}_t,\ \tilde{u}_t,\ \psi_n+\tilde{\varphi}_n)$$

$$-2\delta\tau A(\tilde{u}_t,\tilde{u},\ \psi_n+\tilde{\varphi}_n)+B(\tilde{u},\ \tilde{g}_1)-(2\tilde{u}+m\tau\tilde{u}_t,\ J(U+\delta\tau U_t,\ \tilde{\varphi})$$

$$-R_1(U+\tilde{u},\ V+\tilde{v})+R_1(U,\ V)+R(U+\tilde{u},\ V+\tilde{v})-R(U,\ V)+\tilde{f}_1), \tag{4.4}$$

where

$$B(\tilde{u},\ \tilde{g}_1)=2\nu h^{n-1}\sum_{Q\in\Gamma_h}\tilde{u}(Q)\tilde{g}_1(Q)-\nu h^n\sum_{Q\in\Gamma_h}\tilde{g}_1^2(Q)+2\sigma\nu\tau h^{n-1}\sum_{Q\in\Gamma_h}\tilde{u}(Q)\tilde{g}_{1t}(Q)$$

$$+\nu m\tau h^{n-1}\sum_{Q\in\Gamma_h}\tilde{u}_t(Q)\tilde{g}_1(Q)-\nu\tau h^n\left(\frac{m}{2}+\sigma\right)\sum_{Q\in\Gamma_h}\tilde{g}_1(Q)\tilde{g}_{1t}(Q)$$

$$-\frac{1}{2}\nu m\sigma\tau^2h^n\sum_{Q\in\Gamma_h}\tilde{g}_{1t}^2(Q)+\nu m\sigma\tau^2h^{n-1}\sum_{Q\in\Gamma_h}\tilde{u}_t(Q)\tilde{g}_{1t}(Q).$$

In this section, M_i denotes a constant dependent on U,V and ψ. From Lemma 6, we obtain

$$|\tau(2\delta-m)(\tilde{u}_t,\ J(\tilde{u},\ \psi+\tilde{\varphi}))|\leqslant\delta\tau\|\tilde{u}_t\|^2+\frac{\tau}{4\varepsilon}\|J(\tilde{u},\ \psi+\tilde{\varphi})\|^2$$

$$\leqslant\varepsilon\tau\|\tilde{u}_t\|^2+\frac{\tau M_1}{\varepsilon}(\|\tilde{u}\|_1^2+\|\tilde{u}\|^2+h^{-n}\|\Delta\tilde{\varphi}\|^2\|\tilde{u}\|^2+h^{-n}\|\tilde{\varphi}\|_1^2\|\tilde{u}\|_1^2).$$

Since $p\Delta\tilde{\varphi}=\tilde{u}-\tilde{v}-\tilde{f}_3$ and $(1,\ \tilde{\varphi})=0$, from Lemma 9, the above term is not larger than

$$\varepsilon\tau\|\tilde{u}_t\|^2+\tau M_2(\|\tilde{u}\|^2+\|\tilde{u}\|_1^2)(1+h^{-n}(\|\tilde{u}\|^2+\|\tilde{v}\|^2+\|\tilde{f}_3\|^2+\|\tilde{g}_3\|_{\Gamma_h}^2)). \tag{4.5}$$

From Lemma 7, we have

$$|A(\tilde{u},\tilde{u},\psi_n+\tilde{\varphi}_n)|\leqslant M_3(|\tilde{g}_3|_{\Gamma_h}+|g_3|_{\Gamma_h})\Big\{\|\tilde{u}\|_{\Gamma_h}^2+h^{n-1}\sum_{l=1}^{n}\Big(\sum_{Q\in\Gamma_{lL}}\tilde{u}^2(Q-he_l)$$

$$+\sum_{Q\in\Gamma_{lR}}\tilde{u}^2(Q+he_l)\Big)\Big\}\leqslant M_4(|\tilde{g}_3|_{\Gamma_h}+|g_3|_{\Gamma_h})(\varepsilon\|\tilde{u}\|_1^2+\|\tilde{u}\|^2). \tag{4.6}$$

Similarly we get

$$\left|\frac{\delta m\tau^2}{2}A(\tilde{u}_t,\ \tilde{u}_t,\ \psi_n+\tilde{\varphi}_n)\right|\leqslant M_5(|\tilde{g}_3|_{\Gamma_h}+|g_3|_{\Gamma_h})(\varepsilon\tau^2\|\tilde{u}_t\|_1^2+\tau^2\|\tilde{u}_t\|^2), \tag{4.7}$$

$$|m\tau A(\tilde{u}_t,\ \tilde{u},\ \psi_n+\tilde{\varphi}_n)|\leqslant M_6(|\tilde{g}_3|_{\Gamma_h}+|g_3|_{\Gamma_h})(\varepsilon\tau^2\|\tilde{u}_t\|_1^2+\tau^2\|\tilde{u}_t\|^2+\varepsilon\|\tilde{u}\|_1^2+\|\tilde{u}\|^2). \tag{4.8}$$

From Lemma 7, we have

$$|B(\tilde{u}, \tilde{g}_1)|\leqslant \varepsilon\|\tilde{u}\|_1^2+\varepsilon\tau^2\|\tilde{u}_t\|_1^2+M_7(\|\tilde{u}\|^2+\tau^2\|\tilde{u}_t\|^2+\|\tilde{g}_1\|_{\Gamma_\lambda}^2+\tau^2\|\tilde{g}_{1t}\|_{\Gamma_\lambda}^2). \tag{4.9}$$

From Lemma 9, we have

$$\begin{aligned}|(2\tilde{u}+m\tau\tilde{u}_t, J(U+\delta\tau U_t, \tilde{\varphi}))| &\leqslant \varepsilon\tau\|\tilde{u}_t\|^2+M_8(\|\Delta\tilde{\varphi}\|^2+\|\tilde{\varphi}\|_1^2+\|\tilde{u}\|^2)\\ &\leqslant \varepsilon\tau\|\tilde{u}_t\|^2+M_9(\|\tilde{u}\|^2+\|\tilde{v}\|^2+\|\tilde{f}_3\|^2+\|\tilde{g}_3\|_{\Gamma_\lambda}^2).\end{aligned} \tag{4.10}$$

From Lemma 6, we have

$$\begin{aligned}&|(2\tilde{u}+m\tau\tilde{u}_t, R_1(U+\tilde{u}, V+\tilde{v})-R_1(U, V))|\\ &\quad\leqslant \varepsilon\tau\|\tilde{u}_t\|^2+M_{10}\|\tilde{u}\|^2(h^{-n}\|\tilde{u}\|^2+h^{-n}\|\tilde{v}\|^2+1).\end{aligned} \tag{4.11}$$

We suppose that $U+V>r_0>0$, and $G(\tilde{u}, \tilde{v})$ is such a function that if $\|\tilde{u}\|\leqslant N_1h^{n/2}$, $\|\tilde{v}\|\leqslant N_1h^{n/2}$, N_1 is a suitably small positive constant, then $|G(\tilde{u}, \tilde{v})|$ is bounded. We then have

$$\begin{aligned}&|(2\tilde{u}+m\tau\tilde{u}_t, R(U+\tilde{u}, V+\tilde{v})-R(U, V))|\\ &\quad\leqslant \varepsilon\tau\|\tilde{u}_t\|^2+G(\tilde{u}, \tilde{v})(\|\tilde{u}\|^2+\|\tilde{v}\|^2)(h^{-n}\|\tilde{u}\|^2+h^{-n}\|\tilde{v}\|^2+1).\end{aligned} \tag{4.12}$$

Clearly,

$$|(2\tilde{u}+m\tau\tilde{u}_t, \tilde{f}_1)|\leqslant \varepsilon\tau\|\tilde{u}_t\|^2+M_{11}(\|\tilde{u}\|^2+\|\tilde{f}_1\|^2). \tag{4.13}$$

Now we chooce m. If $\sigma>\frac{1}{2}$, we put

$$m>\max\left(\frac{2\sigma}{2\sigma-1}, 1+m_0\right), \quad m_0>0.$$

Then

$$\tau(m-1)\|\tilde{u}_t\|^2+\nu\tau^2\left(\sigma m-\frac{m}{2}-\sigma\right)\|\tilde{u}_t\|_1^2\geqslant m_0\tau\|\tilde{u}_t\|^2-M_{12}\tau^2h\|\tilde{g}_{1t}\|_{\Gamma_\lambda}^2. \tag{4.14}$$

If $\sigma=\frac{1}{2}$, we put $m>1+2n\lambda\nu+m_0$. Because

$$\tau\|\tilde{u}_t\|_1^2\leqslant 4n\lambda\|\tilde{u}_t\|^2+\tau h\|\tilde{g}_{1t}\|_{\Gamma_\lambda}^2, \tag{4.15}$$

(4.14) holds still. If $\sigma<\frac{1}{2}$, $\lambda<\frac{1}{2n\nu(1-2\sigma)}$, then we put

$$m>\frac{1+4n\lambda\nu\sigma+m_0}{1+4n\lambda\nu\sigma-2n\lambda\nu}.$$

Because of (4.15), we have (4.14) still. Substituting (4.5)—(4.15) into (4.4), we obtain the follwing error estimation for $\|\tilde{u}\|^2$:

$$\begin{aligned}&\|\tilde{u}\|_t^2+\tau(m_0-M_{13}(\varepsilon+\tau)-M_{14}(\varepsilon+\tau)(|\tilde{g}_3|_{\Gamma_\lambda}+|g_3|_{\Gamma_\lambda}))\|\tilde{u}_t\|^2+\nu\|\tilde{u}\|_1^2\\ &\quad+\nu\tau\left(\sigma+\frac{m}{2}\right)(\|\tilde{u}\|_1^2)_t\leqslant \tilde{f}(\tilde{u}, \tilde{v})\|\tilde{u}\|_1^2+(M_{15}+G(\tilde{u}, \tilde{v}))(\|\tilde{u}\|^2+\|\tilde{v}\|^2)\\ &\quad\times(h^{-n}\|\tilde{u}\|^2+h^{-n}\|\tilde{v}\|^2+\tau h^{-n}(\|\tilde{g}_3\|_{\Gamma_\lambda}^2+\|\tilde{f}_3\|^2)+|\tilde{g}_3|_{\Gamma_\lambda}+|g_3|_{\Gamma_\lambda}+1)\\ &\quad+M_{16}(\|\tilde{f}_1\|^2+\|\tilde{f}_3\|^2+\|\tilde{g}_1\|_{\Gamma_\lambda}^2+\tau^2\|\tilde{g}_{1t}\|_{\Gamma_\lambda}^2+\|\tilde{g}_3\|_{\Gamma_\lambda}^2),\end{aligned} \tag{4.16}$$

where

$$\begin{aligned}\tilde{f}(\tilde{u}, \tilde{v})=\{&-\nu+\tau M_2[1+h^{-n}(\|\tilde{u}\|^2+\|\tilde{v}\|^2+\|\tilde{f}_3\|^2+\|\tilde{g}_3\|_{\Gamma_\lambda}^2)]\\ &+\varepsilon M_4(|\tilde{g}_3|_{\Gamma_\lambda}+|g_3|_{\Gamma_\lambda})+M_{17}\varepsilon\}.\end{aligned}$$

We can get a similar estimation for $\|\tilde{v}\|^2$. Hence if $|\tilde{g}_3|_{\Gamma_\lambda}<N_2$, N_2, and ε and τ are suitably small, then we can find such a constant $p_0>0$ that

$$\|\tilde{u}\|_t^2+\|\tilde{v}\|_t^2+\nu\|\tilde{u}\|_1^2+\nu\|\tilde{v}\|_1^2+p_0\tau(\|\tilde{u}_t\|^2+\|\tilde{v}_t\|^2)+\tau\nu\left(\sigma+\frac{m}{2}\right)(\|\tilde{u}\|_1^2+\|\tilde{v}\|_1^2)_t$$
$$\leqslant \tilde{R}(\tilde{u},\tilde{v})+f^*(\tilde{u},\tilde{v})+M_{18}\sum_{i=1}^{3}(\|\tilde{f}_i\|^2+\|\tilde{g}_i\|_{\Gamma_h}^2+\tau^2\|\tilde{g}_{it}\|_{\Gamma_h}^2), \tag{4.17}$$

where
$$f^*(\tilde{u},\tilde{v})=\tilde{f}(\tilde{u},\tilde{v})(\|\tilde{u}\|_1^2+\|\tilde{v}\|_1^2),$$
$$\tilde{R}(\tilde{u},\tilde{v})=(M_{19}+G(\tilde{u},\tilde{v})(\|\tilde{u}\|^2+\|\tilde{v}\|^2))(h^{-n}\|u\|^2+h^{-n}\|v\|^2$$
$$+\tau h^{-n}(\|\tilde{g}_3\|_{\Gamma_h}^2+\|\tilde{f}_3\|^2)+|\tilde{g}_3|_{\Gamma_h}^2+|g_3|_{\Gamma_h}^2+1).$$

By summing the above formula for $t=0, \tau, \cdots, (k-1)\tau$, and letting

$$\tilde{\rho}=M_{20}(\|\tilde{u}(0)\|^2+\|\tilde{v}(0)\|^2+\tau\sum_{j=0}^{k}\left[\sum_{i=1}^{3}(\|\tilde{f}_i(j)\|^2+\|\tilde{g}_i(j)\|_{\Gamma_h}^2)\right],$$

$$\tilde{E}(k)=\|\tilde{u}(k)\|^2+\|\tilde{v}(k)\|^2+p_0\tau^2\sum_{j=0}^{k-1}(\|\tilde{u}_t(j)\|^2+\|\tilde{v}_t(j)\|^2)$$
$$+\nu\tau\sum_{j=0}^{k-1}[\|\tilde{u}(j)\|_1^2+\|\tilde{v}(j)\|_1^2],$$

then we obtain

$$\tilde{E}(k)\leqslant\tilde{\rho}+M_{21}\tau\sum_{j=0}^{k-1}[f^*(\tilde{u}(j),\tilde{v}(j))+\tilde{R}(\tilde{u}(j),\tilde{v}(j))]. \tag{4.18}$$

Finally, using Lemma 10, with

$$\eta(k)=\tilde{E}(k),$$
$$\rho=\tilde{\rho}, \quad m=n,$$
$$f^*(\eta(k))=f^*(\tilde{u}(k),\tilde{v}(k)),$$
$$M(\eta(k))=\tilde{R}(\tilde{u}(k),\tilde{v}(k)),$$

we get the following result.

Theorem 1. *If the following conditions hold.*

(1) $\sigma\geqslant\frac{1}{2}$ *or* $\lambda<\frac{1}{2n\nu(1-2\sigma)}$,

(2) $U+V>r_0>0$,

(3) L *is suitably small,*

(4) N_3, N_4, N_5 *are suitably small positive constants, and*

$$\begin{cases}\tau\|\tilde{f}_3\|^2\leqslant N_3h^n,\\ \tau\|\tilde{g}_3\|_{\Gamma_h}^2\leqslant N_4h^n,\\ |\tilde{g}_3|_{\Gamma_h}\leqslant N_5,\end{cases}$$

then there exist such constants M_1, M_2 *and* N *that for all* $\tilde{\rho}<Nh^n$, $k\tau<T_0$, *we have*

$$\tilde{E}(k)\leqslant M_1\tilde{\rho}e^{M_2k\tau},$$

where T_0 *is a suitably small positive constant dependent on* $\tilde{\rho}$.

Remark 2. If $\tilde{\rho}\to0$ as $h\to0$, the scheme (3.5) is convergent.

References

[1] M. S. Mock, An initial value problem from semiconductor device theory, *SIAM. J. Math. Anal.*, 5 (1974), 597—613.

[2] Kuo Pen-yu, A class of difference scheme for two dimensional viscous vorticity equation, *Acta Mathematics Sinica*, 1974, 242—258.
[3] Kuo Pen-yu, Numerical method of incompressible viscous flow, *Scientia Sinica*, 1977, 287—304.
[4] Kuo Pen-yu, Difference method of fluid dynamics-Numerical solution of Primitive equation, *Scientia Sinica*, 1981, 297—312.
[5] Kuo Pen-yu, Resolution numérique de fluide compressible, C. R. Acad. Sc. Paris, t 291 (1980), Serie A, 167—170.
[6] Kuo Pen-yu, Numerical solution of an initial-value problem for a semiconductor device, *COMPEL*, **2** (1983), 57—75.