

# FDM solvers for 1D wave problem (II)

huangyf15

## 目录

<b>2 Problem 2: 差分格式的精度</b>	<b>4</b>
2.1 差分格式的精度分析	4
2.2 单调格式的精度局限	5
<b>3 Problem 3: 差分格式的色散耗散特性</b>	<b>6</b>
3.1 基础格式的构造	6
3.2 高分辨率格式的构造	7
3.3 色散耗散特性初步分析	8
3.3.1 程序实现与初步测试结果	8
3.3.2 基础格式与优化后格式的参数	9
3.3.3 色散耗散特性综合分析	10
3.4 算例验证与误差分析	12
3.4.1 间断算例的验证	12
3.4.2 多尺度算例的验证	12
3.4.3 高波数成分对格式分辨的影响	12
3.4.4 误差分析与改进思路	12

## 2 Problem 2: 差分格式的精度

已知一维线性波动方程

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0$$

有如下形式的差分格式

$$u_j^{n+1} = \sum_{k=-l}^l b_k u_{j+k}^n. \quad (2.1)$$

试对上述差分格式的精度进行分析，并回答如下问题：

- 推导其具有  $p$  阶精度的充要条件，并给出  $p$  可取到的最大值；
- 当  $b_k \geq 0$  ( $\forall k$ ) 时称上述格式为单调格式，证明上述方程的单调格式至多到一阶精度。

### 2.1 差分格式的精度分析

由题意，时间方向采用前差格式，从而可写

$$u_j^{n+1} = u_j^n + c \Delta x \frac{\partial u_j^n}{\partial x}, \quad (2.2)$$

其中  $c := \frac{a \Delta t}{\Delta x}$  为 CFL 数。

下面将  $u_j^{n+1}$  在  $(j, n)$  附近作 Taylor 展开，有

$$\begin{aligned} \tilde{u}_j^{n+1} &= \sum_{k=-l}^l b_k u_{j+k}^n \\ &= \sum_{k=-l}^l b_k E_x^k u_j^n \\ &= \sum_{k=-l}^l b_k e^{k \Delta x D_x} u_j^n \\ &= \sum_{k=-l}^l b_k \sum_{m=0}^{\infty} \frac{(k \Delta x D_x)^m}{m!} u_j^n \\ &= \sum_{m=0}^{\infty} \left[ \sum_{k=-l}^l \frac{k^m}{m!} b_k \right] (\Delta x D_x)^m u_j^n. \end{aligned} \quad (2.3)$$

联立方程 (2.2) 和 (2.3)，考虑到  $L.T.E. := \tilde{u}_j^{n+1} - u_j^{n+1}$ ，格式具有一阶精度等价于

$$\begin{aligned} m=0 : \quad & \sum_{k=-l}^l b_k = 1, \\ m=1 : \quad & \sum_{k=-l}^l k b_k = c. \end{aligned}$$

达到二阶精度则还需补充

$$m=2 : \quad \sum_{k=-l}^l k^2 b_k = 0.$$

一般地, 上述格式具有  $p$  阶空间精度的充要条件是

$$\begin{cases} \sum_{k=-l}^l b_k = 1, \\ \sum_{k=-l}^l k b_k = c, \\ \sum_{k=-l}^l k^2 b_k = 0, \\ \vdots \\ \sum_{k=-l}^l k^p b_k = 0. \end{cases} \quad (2.4)$$

我们断言, 使上述方程恒有解的  $p$  值最大为  $2l$ , 这也可从待定系数  $b_k$  的个数推断出。事实上, 当  $p = 2l$  时, 上述线性方程组的系数矩阵是范德蒙矩阵, 行列式的值为  $\prod_{0 \leq j < i \leq 2l} (x_i - x_j) \neq 0$ , 因此方程组有以  $c$  为参数的唯一解。而当  $p > 2l$  时, 为使超定方程组有解, 必须使右端向量中的参数  $c$  取特定值, 这显然不具有任何普适性, 因此不符合要求。

## 2.2 单调格式的精度局限

显然对于单调格式来说, 一阶精度是可以实现的, 我们只需证明单调格式无法达到二阶精度即可。事实上, 由于对于单调格式成立  $b_k \geq 0$  ( $\forall k$ ), 若还有  $\sum_{k=-l}^l k^2 b_k = 0$ , 则只能有  $b_j = \delta_{j0}$ , ( $j = -l, -l+1, \dots, 0, l-1, l$ ), 从而  $u_j^{n+1} = u_j^n$ , 该格式显然不具有二阶精度, 矛盾! 这样就完成了证明。

### 3 Problem 3: 差分格式的色散耗散特性

要求设计求解  $\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0$  ( $a > 0$ ) 的高分辨率半离散格式。具体地说, 首先设计基础格式使之在均匀网格的给定模版点上达到最高的空间精度, 然后通过降低精度、增加自由参数的方法进行优化。为了明确优化思路, 我们首先对优化依据、算例构造与误差分析这几个问题进行如下的简要讨论。

**优化依据** 即要求色散参数在某种范数意义上下取极值。题干主要叙述了以调节参数  $k_c$  为基础的积分限控制以及以调节参数  $\nu$  为基础的权函数控制两种方法; 除此之外, 还可以采用连续意义下的  $p$  范数、离散意义下的  $\ell^p$  范数及其截断形式等。另一方面, 在优化格式的色散特性时往往要牺牲其耗散特性, 其最基本的要求是正耗散, 这等价于一个不等式约束。对此我们可以采取罚函数法或后验估计的方法, 也可以采用文献 [1] 中的方法将色散耗散参数相互解耦。这里我们采用权函数控制的方法, 格式的构造与色散耗散特性分析等详细讨论详见第 3.2 和 3.3 节。

**算例构造** 应用上述格式具体算例在相同网格上计算相同时间。作为第一个算例, 也是最主要的一个算例, 多尺度算例的初值可采用如式 (3.1) 的形式

$$u(x, 0) = 1 + \varepsilon \sum_{k=1}^{N/4} \left( \sqrt{E(k)} \sin(2\pi k(x + \psi_k)) \right), \quad x \in [0, 1]. \quad (3.1)$$

其中,  $\varepsilon = 0.1$ ,  $E(k) = \left(\frac{k}{k_0}\right)^4 e^{-2(k/k_0)^2}$ ,  $\psi_k$  是  $0 \sim 1$  间的随机数,  $N$  是网格数 (可取 128 或 256), 采用周期性边界条件, 详见第 3.4.2 节。算例中的参数  $k_0$  越大, 高波数的成分越丰富, 格式的分辨也就越困难, 下面将取  $5 \sim 20$  间的不同整数来考察其影响, 详见第 3.4.3 节。作为第二个算例, 我们还将构造一个具有一定间断的简单算例作为色散耗散特性分析的补充, 详见第 3.4.1 节。

**误差分析** 对比与精确解的误差大小时, 可采用 1 范数进行度量, 即

$$\|U^n - U_e^n\| := \frac{1}{N} \sum_j |u_j^n - u_{e,j}^n|. \quad (3.2)$$

下面将要看到, 这种度量误差的方式实际并不是最优的, 需要进行一定的色散补偿, 有关这一问题的详细讨论将在第 3.4.4 小节展开。

#### 3.1 基础格式的构造

这里构造基础格式的方法与第 2 大节是类似的。使用如下的半离散格式

$$\frac{\partial u_j^n}{\partial t} = -a \frac{\partial u_j^n}{\partial x}.$$

其中, 空间方向的离散记为

$$\frac{\partial u_j^n}{\partial x} \simeq: f(u_{j-l}^n, \dots, u_{j+r}^n) := \sum_{k=-l}^r b_k u_{j+k}^n. \quad (3.3)$$

为保证足够的格式整体精度，时间方向采用 4 阶 Runge-Kutta 方法进行离散

$$\begin{cases} u_j^{(1)} = u_j^n - \frac{c}{4}f(u^n), \\ u_j^{(2)} = u_j^n - \frac{c}{3}f(u^{(1)}), \\ u_j^{(3)} = u_j^n - \frac{c}{2}f(u^{(2)}), \\ u_j^{n+1} = u_j^n - \frac{c}{1}f(u^{(3)}). \end{cases} \quad (3.4)$$

其中， $c := \frac{a\Delta t}{h}$  为 CFL 数。

对式 (3.3) 采用不同的差分近似即可得到不同精度的格式。与第 2 大节类似地可以得到，格式具有  $p$  阶空间精度的充要条件是

$$\begin{cases} \sum_{k=-l}^r b_k = 0, \\ \sum_{k=-l}^r kb_k = 1, \\ \sum_{k=-l}^r k^2b_k = 0, \\ \vdots \\ \sum_{k=-l}^r k^pb_k = 0. \end{cases} \quad (3.5)$$

这里，给定模板点可达到的最高阶精度为  $p = l + r$ 。为了下一小节优化的方便，我们将其写成矩阵形式，即  $Kb = f$ 。上述格式的色散与耗散误差可以分别表示为

$$\begin{aligned} \Re(k') &= \sum_{j=-l}^r b_j \sin(jk), \\ \Im(k') &= - \sum_{j=-l}^r b_j \cos(jk). \end{aligned}$$

### 3.2 高分辨率格式的构造

我们仍采用式 (3.3) 形式的格式，通过降低精度要求、增加自由参数的方式设法提高格式的分辨率。这里我们只对于模板点填满  $-l \sim r$  的情形进行讨论，并假设左右各取  $\text{prL}$ ,  $\text{prR}$  个模板点的系数作为自由参数。这等价于减少了  $\text{prL} + \text{prR}$  个精度约束，使得格式损失了  $\text{prL} + \text{prR}$  阶精度。考虑到格式具有  $p^* := l + r - \text{prL} - \text{prR}$  阶精度意味着  $Kb = f$  的前  $p^* + 1$  行成立，因此先将上述方程改写为

$$\begin{pmatrix} K_{md11} & K_{mm} & K_{md12} \\ K_{md21} & K_{dm} & K_{md22} \end{pmatrix} \begin{pmatrix} b_{d1} \\ b_m \\ b_{d2} \end{pmatrix} = \begin{pmatrix} f_m \\ f_d \end{pmatrix}$$

其中，

$$K_{md11} \in \mathbb{R}_{p^* \times \text{prL}},$$

$$K_{md12} \in \mathbb{R}_{p^* \times \text{prR}},$$

$$K_{mm} \in \mathbb{R}_{p^* \times p^*}.$$

进而可写为

$$\begin{pmatrix} K_{mm} & K_{md1} \\ K_{dm} & K_{md2} \end{pmatrix} \begin{pmatrix} b_m \\ b_d \end{pmatrix} = \begin{pmatrix} f_m \\ f_d \end{pmatrix}$$

其中,

$$K_{mdi} = (K_{mdi1} \ K_{mdi2}), \\ b_d = (b_{d1} \ b_{d2}).$$

由于我们放宽了精度上的要求, 因此只需满足前一部分等式, 即只要求

$$K_{mm}b_m + K_{md1}b_d = f_m \quad (3.6)$$

解得

$$b_m = K_{mm}^{-1}(f_m - K_{md1}b_d). \quad (3.7)$$

这里,  $b_d$  是作为待定参数出现的。

上述格式的色散和耗散误差可以写为

$$\Re(k') = \sum_{j=-l}^r b'_j \sin(jk), \\ \Im(k') = - \sum_{j=-l}^r b'_j \cos(jk).$$

只要使得  $\Re(k')$  和  $k$  在某种意义上最为接近 (优先低波数范围) 同时保证格式具有正耗散特征, 通过极值条件求得待定参数  $b_d$ , 再带入式 (3.5) 即可求得  $b_m$ , 从而确定了所有的  $b'_i$  ( $i = -l, \dots, r$ )。这里我们采用极小化加权积分的形式, 即取  $b_d$  使得加权积分

$$E(b_d) = \int_0^\pi e^{-\nu k} (\Re(k') - k)^2 dk \quad (3.8)$$

达到最小值, 这等价于

$$\frac{\partial E}{\partial b_{d_i}} = 0, i = 1, \dots, r, r+1, \dots, r+l.$$

我们指出, 上述操作实际上等价于求取函数  $E(b_{-l}, \dots, b_r)$  在约束 (3.6) 下的条件极值问题。一个重要的推论是, 无论放开哪一部分参数 (这里不要求放开的参数在模板点集合的外缘), 由于上述条件极值的约束是相同的, 因此只要放开的参数个数相同并且相应的 *Jacobi* 矩阵非奇异, 那么最终求得的格式形式应该是相同的。

### 3.3 色散耗散特性初步分析

#### 3.3.1 程序实现与初步测试结果

我们使用 MATLAB 编程, 通过符号运算实现了任意左右模板点数  $l, r$  ( $l \geq r$ ) 以及任意自由参数个数  $\text{prL}, \text{prR}$  ( $l - \text{prL} \geq r - \text{prR}$ ) 下的差分近似格式参数的计算, 同时实现了对应格式的色散耗散特性分析, 现将细节叙述如下。

- 色散特性的优化方案采用的是极小化加权积分方案, 并取参数  $\nu = 10$ , 这样可以在保证低波数结构计算的准确性的基础上充分考虑高波数的影响;
- 色散耗散特性的表征方法有两种, 一是色散耗散特性曲线, 较为粗略; 二是对应于某一相对误差  $\varepsilon = 0.01$  或  $0.05$  的波数最大值  $k^*$ , 较为精确。

为了此后叙述的方便起见，我们将通过差分方法得到的格式统一命名为

$$U_{l[i]}D_{r[k]}[F]_{[M][LmRn]}.$$

其中，带 F (FULL) 时该格式的模板点填满  $-l \sim r$ ， $l, r$  代表上、下游终止模板点的标号；不带 F 时该格式的模板点由  $l_i, r_j$  单独指定；带  $_M$  (AVERAGE) 时代表该格式是通过其他格式的某种平均的方式得到；带  $_{LmRn}$  的为将上、下游分别放开  $l, r$  个自由度后获得的色散优化格式。

作为程序的测试算例，下面我们给出了 U14D13F 的格式参数 **b**，图 1 则给出了该格式的色散耗散特性曲线。

(0.000000003570573	-0.000000103839268	0.000001462696994	-0.000013300384324 ...
0.000087807204685	-0.000448944950139	0.001852674289017	-0.006355386902595 ...
0.018549276635815	-0.047034867839653	0.105965156215500	-0.218763400180497 ...
0.438716811757672	-1.021057201523688	0.105338339112007	0.846590829614485 ...
-0.320658008922075	0.138791770734302	-0.057963376626504	0.021998649957349 ...
-0.007340678509788	0.002098962915800	-0.000501183495908	0.000096891651182 ...
-0.000014539123202	0.000001586567302	-0.000000111901657	0.000000003826518)

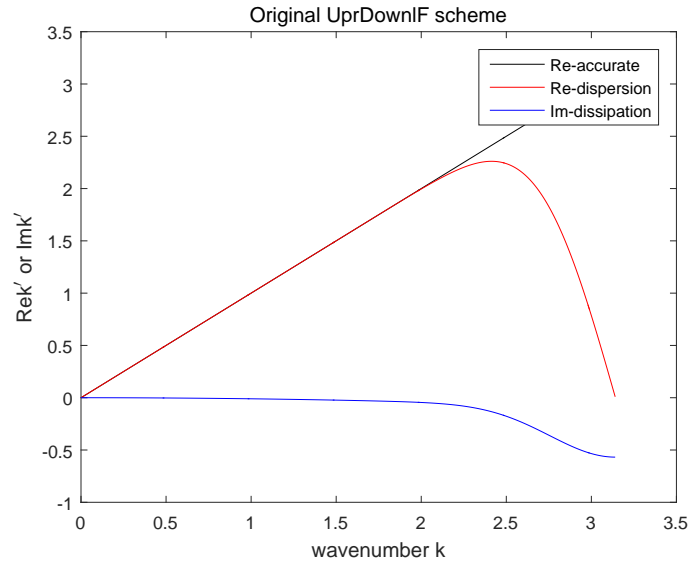


图 1: Example (U14D13F)

### 3.3.2 基础格式与优化后格式的参数

为了格式优化过程的典型性起见，我们选取  $l = 4, r = 3$  对应的格式作为基础格式。在对格式进行优化时，我们重点考察如下两种情形：prL = 0 且 prR = 1, 2, 3；prL = prR = 1 或 2。

下面给出上述格式的各模板参数, 为方便起见将  $\text{prL} = m, \text{prR} = n$  的格式记为  $\text{LmRn}$ 。

$$\begin{aligned}
 b_{\text{L0R0}} &= \begin{pmatrix} 0.007142857142857 & -0.066666666666667 & 0.300000000000000 & -1.000000000000000 & \dots \\ 0.250000000000000 & 0.600000000000000 & -0.100000000000000 & 0.009523809523810 & \dots \end{pmatrix} \\
 b_{\text{L0R1}} &= \begin{pmatrix} 0.016643320975409 & -0.133169913494530 & 0.499509740483590 & -1.332516234139316 & \dots \\ 0.582516234139316 & 0.400490259516410 & -0.033496753172137 & 0.000023345691258 & \dots \end{pmatrix} \\
 b_{\text{L0R2}} &= \begin{pmatrix} 0.023991716885747 & -0.182924576920408 & 0.643721406921766 & -1.564448471803848 & \dots \\ 0.806027932071418 & 0.271435564596607 & 0.007837370521312 & -0.005640942272594 & \dots \end{pmatrix} \\
 b_{\text{L0R3}} &= \begin{pmatrix} 0.092498441381782 & -0.446232303982341 & 1.078202135187173 & -2.184971564563121 & \dots \\ 1.811732204331870 & -0.856371286770921 & 0.656326277084423 & -0.151183902668865 & \dots \end{pmatrix} \\
 b_{\text{L1R1}} &= \begin{pmatrix} -0.050875476965703 & 0.153297614279266 & 0.198599330604171 & -1.761819173410236 & \dots \\ 1.942639462480186 & -0.974075850930082 & 0.610856008124017 & -0.118621914181620 & \dots \end{pmatrix} \\
 b_{\text{L2R2}} &= \begin{pmatrix} 0.042825431460906 & -0.175044011963002 & 0.016486627296339 & -0.123542651699957 & \dots \\ 0.497961138631428 & -1.691167659891931 & 2.050675425759381 & -0.618194299593164 & \dots \end{pmatrix}
 \end{aligned}$$

### 3.3.3 色散耗散特性综合分析

我们首先给出上述格式的色散耗散特性曲线, 如图 2 ~ 3 所示。

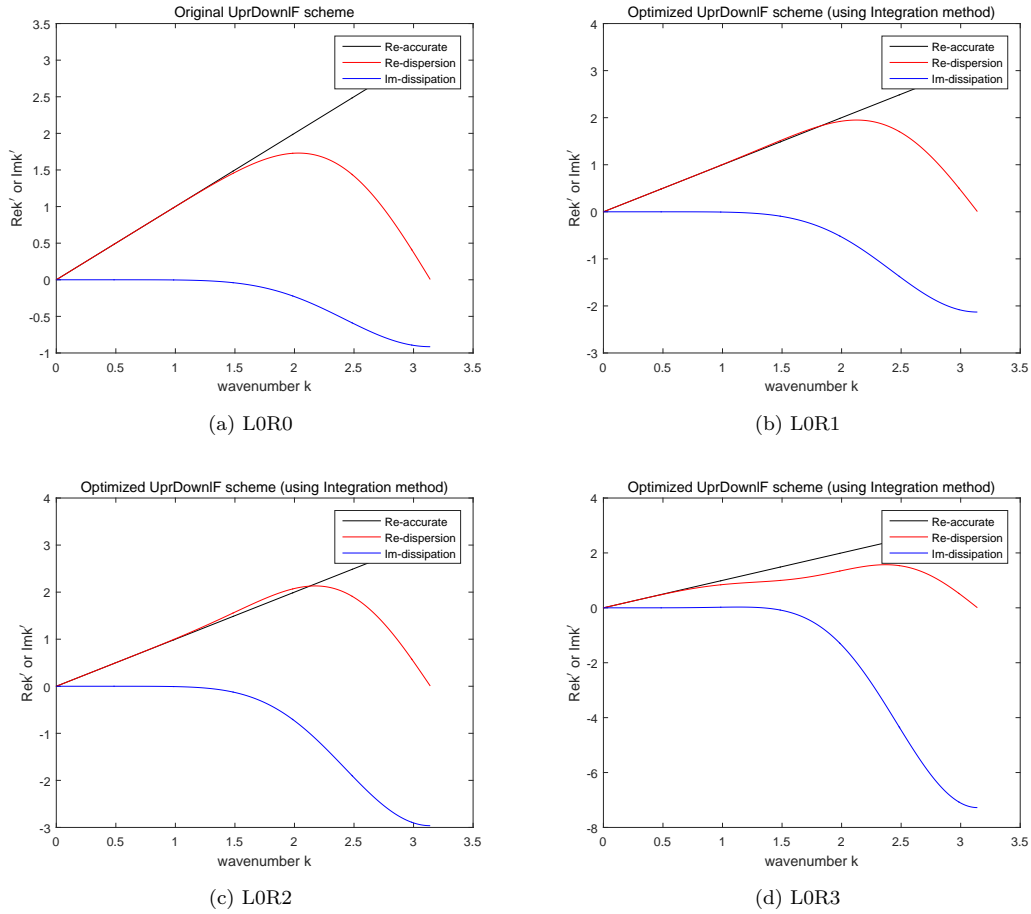


图 2: LORX scheme



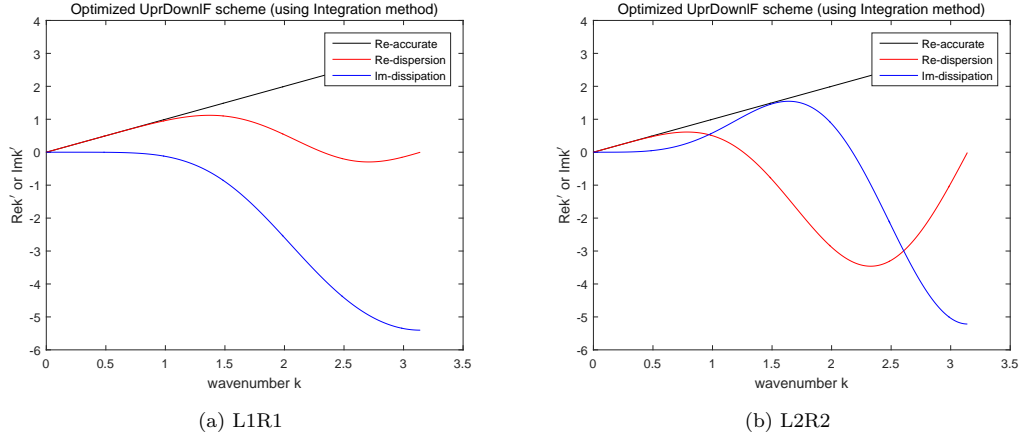


图 3: LmRm scheme

特定格式及给定相对误差限下的波数分界值  $k^*$  列举如下。

$$\begin{aligned}
 k_{L0R0,0.05}^* &= 1.6900000000000001, & k_{L0R0,0.01}^* &= 1.3300000000000001; \\
 k_{L0R1,0.05}^* &= 2.0400000000000001, & k_{L0R1,0.01}^* &= 1.1700000000000001; \\
 k_{L0R2,0.05}^* &= 1.5000000000000001, & k_{L1R1,0.05}^* &= 1.0400000000000001; \\
 k_{L0R3,0.05}^* &= 0.6900000000000000, & k_{L2R2,0.05}^* &= 0.5200000000000000.
 \end{aligned}$$

根据上述结果，我们可作如下分析：

1. 由图 2a ~ 2b 可见，当在基础格式的基础上放开 1 个自由参数时，格式的色散特性有较好的优化。从波数分界点值  $k^*$  的比较中也可进一步看出，尽管在相对误差限  $\varepsilon$  为 0.01 时仍是基础格式的表现较好，但当  $\varepsilon = 0.05$  时优化后格式的表现便更胜一筹，这说明优化过程确实提升了平均意义上的色散特性。
2. 由图 2c ~ 2d 以及图 3a ~ 3b 可见，当在基础格式的基础上放开 2 个及以上自由度时，色散曲线显示此时会呈现一定的负优化特征，色散特性曲线将更加偏离真实值。特别地，当放开上游节点的自由度时，格式甚至可能出现负耗散现象，导致格式稳定性的丧失。然而，按照第 3.2 节的讨论，色散优化格式参数应该仅与放开模板点的个数有关，而与其位置无关；上述计算结果的出现很有可能是 MATLAB 在矩阵求逆时损失精度造成的。这是由于在针对实部进行色散特性优化的过程中，并未考虑虚部耗散项的影响，因此有可能牺牲格式的耗散特性。
3. 由图 2 和图 3 可以看到，当格式的色散特性得到某种程度的优化时，其耗散性将明显增强。并且，放开的自由度数越多，耗散增强得越明显。这是由于此处的色散优化是以格式精度的降低为代价的，而格式精度的降低必然导致耗散效应的增强，这主要包括耗散与色散的相对强弱以及低阶耗散项本身的影响。
4. 更进一步说，无论是  $k_c$  方案还是  $\nu$  方案，其对于色散特性的优化思路都具有很大的人为性，尤其是在优化的标准上存在着很大差异。可以设想，一种更为合理的优化标准是将耗散特性作为先验的约束而非后验的验证措施，同时对于色散特性和耗散特性作更为综合的考量。

由此，我们可以得出经验性的结论，即为了取得较好的色散耗散综合优化效果，应该使基础格式具有尽可能高的精度，同时只放开少量下游节点的自由度（尽管这在目前仅是一个归纳式的结果），以保证在格式的色散特性得到充分优化的基础上，其精度和稳定性不会受到较大损失。

### 3.4 算例验证与误差分析

下面我们通过具体的算例对格式的色散耗散特性进行验证与分析。这里，我们选取了两种算例，分别为间断算例和多尺度算例，二者除初值条件外没有其他差别。具体来说，我们令线性波动方程中的波速  $a = 1$ ，选取  $[0, 1]$  作为计算域并将其离散为  $N_x = 256$  的网格；取  $CFL = 0.1$  用于确定时间方向的推进步长，同时尽可能保证格式的稳定性；下面所有算例中截取的时间片为求解域宽度即波长的整数倍，从而确保时间片处的精确解与初始分布相同。

#### 3.4.1 间断算例的验证

图 4 给出了各种格式计算间断算例的结果，下述结果与此前作业以及上一节中的论述是相符的。由于篇幅所限，这里不作为重点展开论述。

#### 3.4.2 多尺度算例的验证

图 5 给出了各种格式计算多尺度算例的结果，下述结果与此前作业以及上一节中的论述是相符的。由于篇幅所限，这里亦不作为重点展开论述。

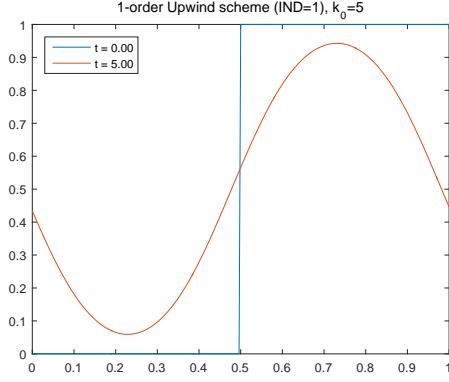
#### 3.4.3 高波数成分对格式分辨的影响

图 6 给出了 U3D2F 格式在计算具有不同波数分布的初值算例时的结果，这里为了突出高波数难以分辨的特点而特地选取了精度与分辨率较低的格式。由图 6a 到图 6d 的演变容易看出，初值所含有的高波数成分越多，格式就越难对其进行分辨。

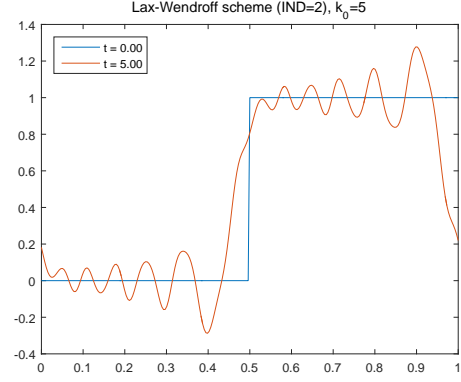
#### 3.4.4 误差分析与改进思路

图 7 给出了各种格式计算多尺度算例时函数值的 1 范数误差，其中格式的编号 IND 见图 5 子图 5a ~ 5g 中的小标题。图中星形标记的误差是严格按照式 (3.2) 计算、未经修正的误差计算值，容易看到它与 5 的计算结果是不相符合的，因此我们考虑通过色散补偿的方式修正误差计算方法。

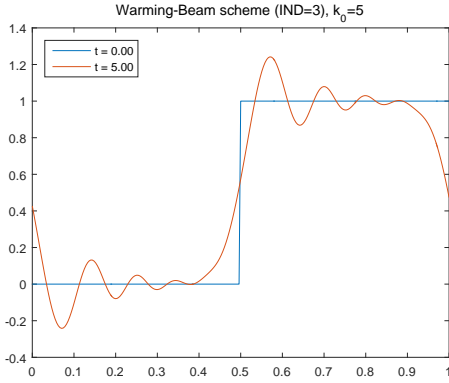
具体来说，就是将色散带来的相对于非定常精确解的偏移量纳入考虑范围，即将式 (3.2) 中的  $u_j^n - u_{e,j}^n$  一项修正为  $u_{j-(T_0/dt)mv}^n - u_{e,j}^n$ ， $mv$  代表了单位时间内的色散偏移量。从图 7 中十字花标记的误差可以看出，修正后的误差计算值与格式的实际分辨率是更相符合的。同时可以看到，进行色散优化后的格式 IND=12 在这种意义下的误差要比优化前的 IND=11 小一些。



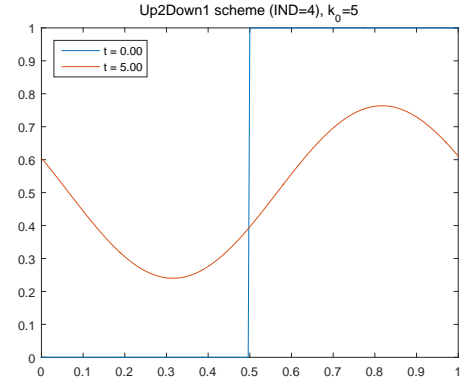
(a) 1-order Upwind (U1D0F)



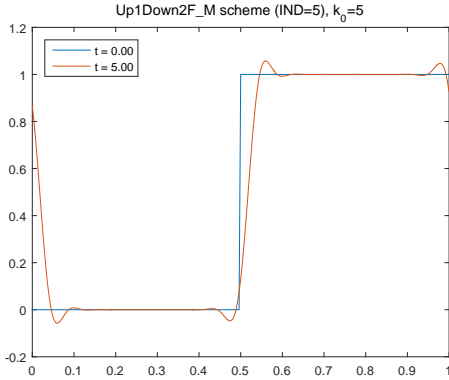
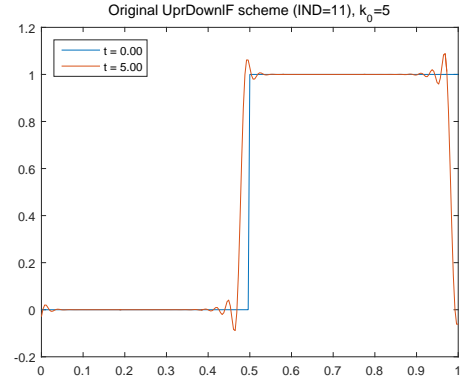
(b) Lax-Wendroff (U1D1F)



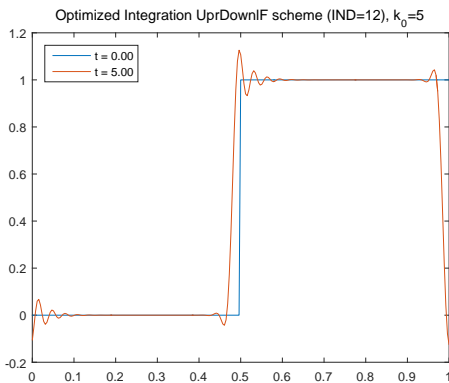
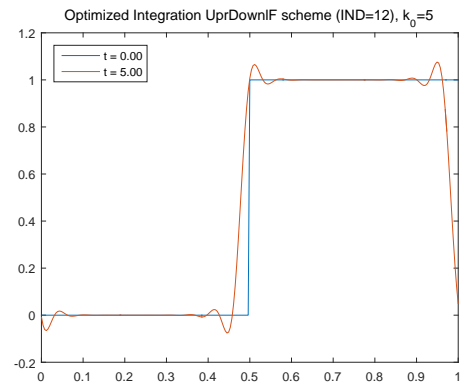
(c) Warming-Beam (U2D0F)

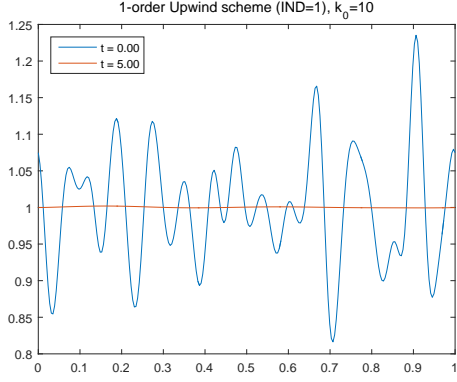


(d) U2D1

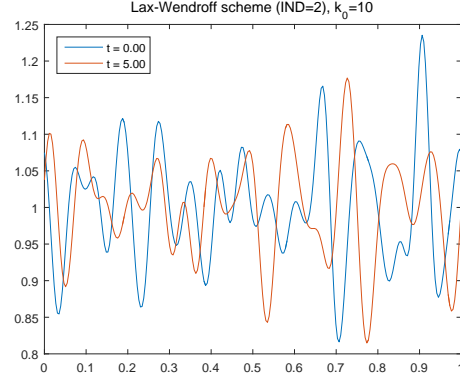
(e) U2D1F<sub>M</sub>

(f) U4D3F

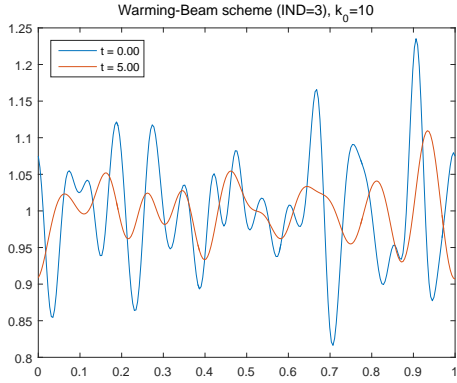
(g) U4D3F<sub>L0R1</sub>(h) U4D3F<sub>L1R1</sub>图 4: Results of U/D<sub>r</sub> schemes on discontinuous examples



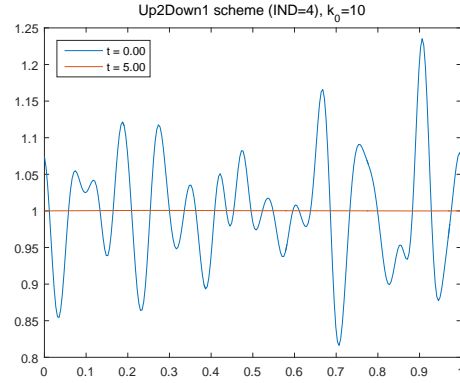
(a) 1-order Upwind (U1D0F)



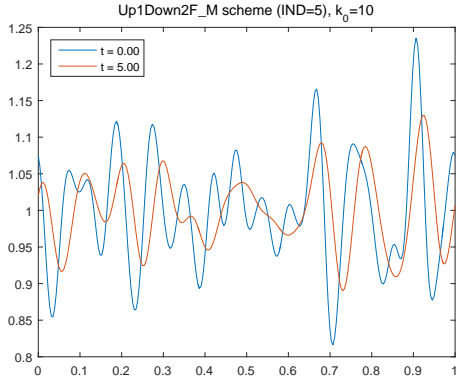
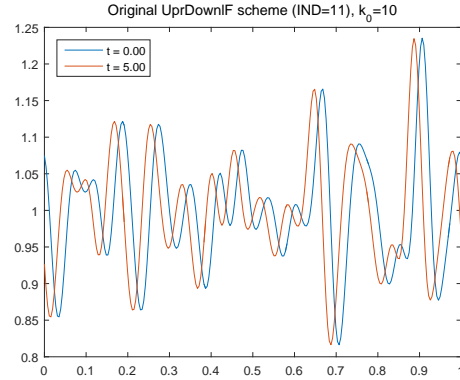
(b) Lax-Wendroff (U1D1F)



(c) Warming-Beam (U2D0F)



(d) U2D1

(e) U2D1F<sub>M</sub>

(f) U3D4F

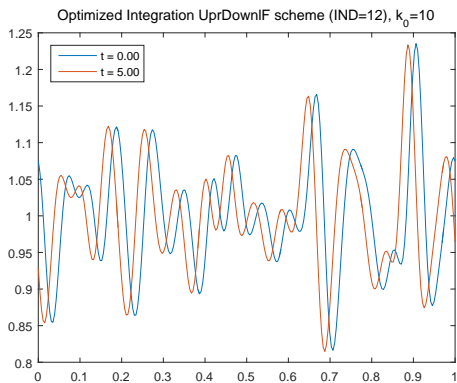
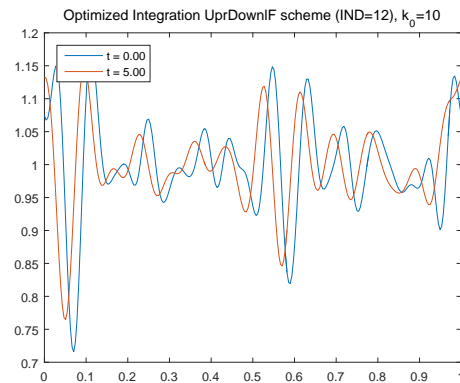
(g) U3D4F<sub>L0R1</sub>(h) U3D4F<sub>L1R1</sub>

图 5: Results of UIdr schemes on multiscale examples

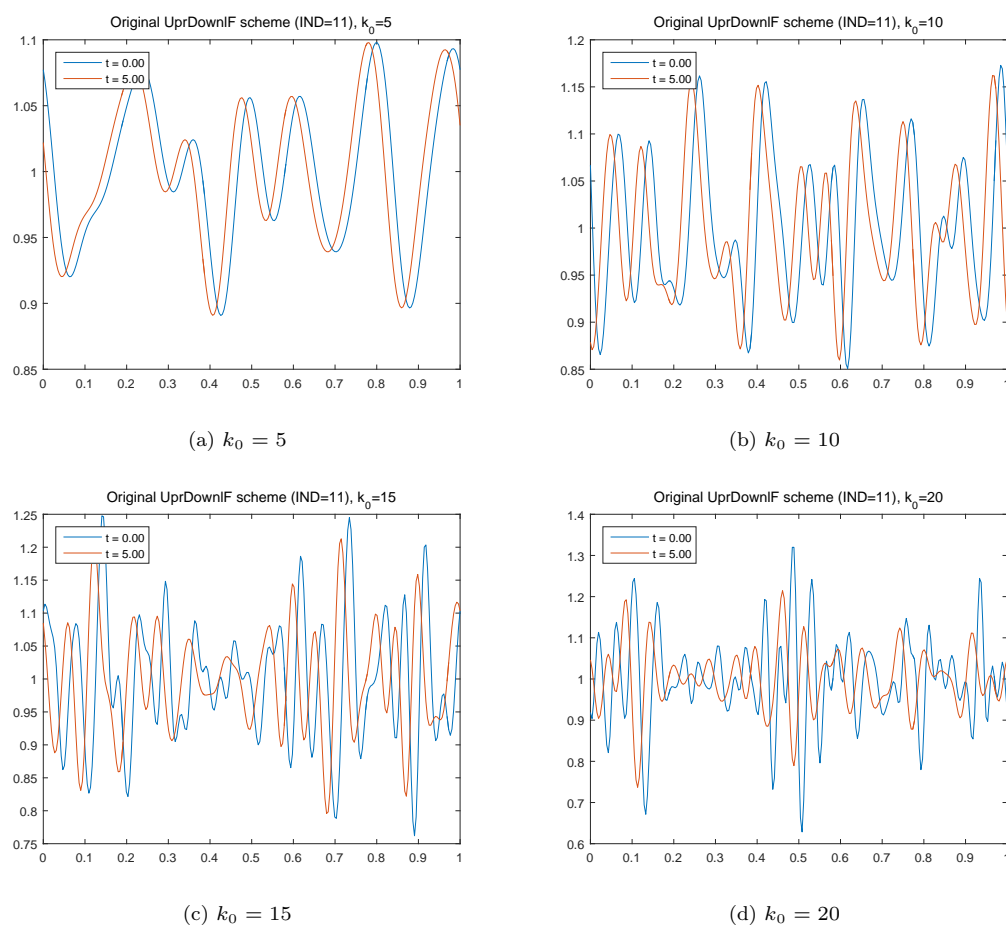


图 6: Results of U3D2F scheme under different wavenumbers

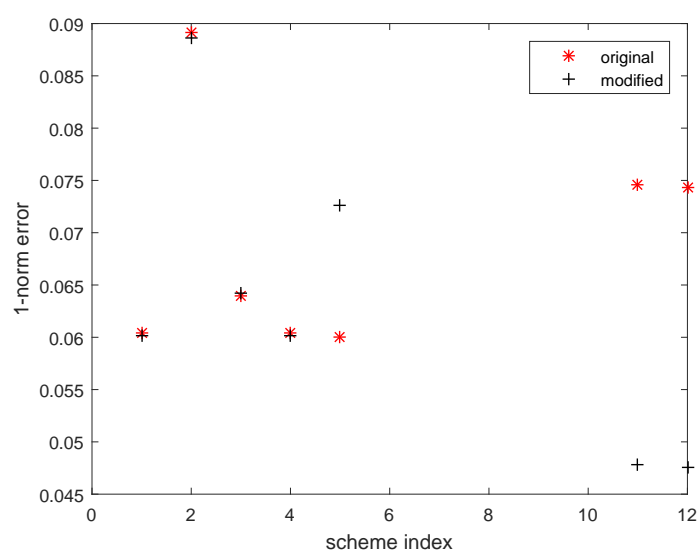


图 7: Error estimation on different schemes

## 参考文献

- [1] Z.-S. Sun et al. Journal of Computational Physics 230 (2011) 4616-4635.
- [2] 任玉新, 陈海昕. 计算流体力学基础. 北京: 清华大学出版社, 2006.
- [3] 陶文铨. 数值传热学 (第 2 版). 北京: 清华大学出版社, 2002.