

学号 20161413

密级 无

东北大学本科毕业论文

基于三维铌酸锂非线性光子晶体的 非线性全息设计与研究

学院名称：理学院

专业名称：应用物理学

学生姓名：谢尘竹

指导教师：高志达 教授

二〇二〇年六月

Nonlinear Holographic Design and Research Based on Three Dimensional Lithium Niobate Nonlinear Photonic Crystals

By Chenzhu Xie

Supervisor: Professor Zhida Gao

Northeastern University

June 2020

郑重声明

本人呈交的学位论文，是在导师的指导下，独立进行研究工作所取得的成果，所有数据、图片资料真实可靠。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确的方式标明。本学位论文的知识产权归属于培养单位。

本人签名：谢尘竹 日期：2020年6月10日

摘要

非线性光子晶体具有空间依赖的二阶非线性系数分布，人们根据各种位相匹配机制设计二阶非线性系数分布，从而有效地操控各种非线性光学效应的发生，铌酸锂(LiNbO_3)晶体具有出色的非线性光学性质，是当前最受欢迎的人工非线性光子晶体材料之一，一维和二维的铌酸锂非线性光子晶体已被广泛地应用于激光变频、非线性光场调控，以及非线性光学成像等领域。然而三维非线性光子晶体由于受到加工技术的限制，在实验上的实现成为近年来非线性光学领域的最大挑战之一。2018年，基于飞秒激光直写技术擦除二阶非线性系数的三维铌酸锂非线性光子晶体横空出世，三维铌酸锂非线性光子晶体拥有更高的空间维度，包含更加丰富的倒格矢，进而允许通过设计三维非线性光子晶体的结构以全方位操控非线性过程的发生，大量基于三维非线性光子晶体的非线性光学研究将进一步开展和进行。

本课题面向飞秒激光直写的三维铌酸锂非线性光子晶体，研究了三维非线性结构中的准相位匹配机制，提出三维 Gerchberg–Saxton (GS) 算法，对于给定三维倒格矢分布，逆向求解三维纯相位体全息结构，接着利用迂回相位编码方法，获得最终的二阶非线性系数三维分布，可通过飞秒激光直写的方式，选择性地对相应位置的二阶非线性系数进行部分擦除，将三维体全息结构“刻”在非线性晶体内。本文利用数值方法模拟了基频光入射下，三维非线性光子晶体内同时实现二次谐波的产生以及高效调控，出射晶体后的二次谐波成功在远场成像。

关键词：铌酸锂；飞秒激光；非线性光子晶体；准相位匹配；非线性全息

ABSTRACT

A nonlinear photonic crystal (NPC) possesses space-dependent second-order nonlinear coefficients, which can effectively control nonlinear optical interactions through quasi-phase matching. Lithium niobate (LiNbO_3) crystal is one of the most popular materials from which to fabricate NPC structures because of its excellent nonlinear optical properties. One- and two-dimensional LiNbO_3 NPCs have been widely utilized in laser frequency conversion, spatial light modulation and nonlinear optical imaging. However, limited by traditional poling methods, the experimental realization of three-dimensional (3D) NPCs remains one of the greatest challenges in the field of nonlinear optics in recent years. In 2008, An experimental demonstration of a 3D LiNbO_3 NPC by using a femto-second laser to selectively erase the nonlinear coefficients in a LiNbO_3 crystal was presented. The effective conversion efficiency is comparable to that of typical quasi-phase-matching processes. Such a 3D LiNbO_3 NPC provides a promising platform for future nonlinear optical studies based on its unique ability to control nonlinear interacting waves in 3D configuration.

This study is based on the three-dimensional lithium niobate nonlinear crystals modulated by femtosecond laser writing system. Through the three-dimensional GS algorithm and adopting certain phase and amplitude constraints, the target object light distribution is repeatedly Fourier transformed into a phase only hologram. Then, according to the three-dimensional Fourier computed volume hologram generated by the three-dimensional GS algorithm, the three-dimensional detoured phase coding method is used to shift the different coordinate positions of focusing beam with respect to the periodic cell structure in the modulation region of the lithium niobate crystal to selectively erase the second-order nonlinear coefficients at the corresponding positions, and "carve" the hologram in the modulation area of the three-dimensional photonic crystal. Finally, under the incidence of the fundamental-frequency light, the second harmonic wave is generated in the crystal and the second harmonic wave after exiting the crystal is imaged in the far field at the same time.

Key words: LiNbO_3 ; Femtosecond laser; Nonlinear photonic crystal; Quasi phase

matching; Nonlinear holography

目 录

郑重声明	I
摘要	II
ABSTRACT	III
1 绪论	1
1.1 引言	1
1.1.1 非线性光学	1
1.1.2 非线性光子晶体	1
1.2 非线性全息	3
1.2.1 非线性全息的三个特征	4
1.2.2 非线性全息的应用	6
1.3 非线性全息编码成像	6
1.3.1 二维非线性菲涅尔全息编码成像	8
1.3.2 三维非线性傅里叶全息编码成像	9
2 三维非线性傅里叶全息成像原理	11
2.1 非线性光学	11
2.1.1 从麦氏方程组的微分形式到非线性电介质中的波动方程	13
2.1.2 磁各向同性的非线性电介质中的波动方程及其解	19
2.2 一维非线性光子晶体中的准相位匹配机制	35
2.2.1 准相位匹配是除完美相位匹配方法外提高转换效率的较好方法	37
2.3 利用三维非线性光子晶体实现傅里叶全息成像	41
2.3.1 三维光子晶体中的准相位匹配机制	42
2.3.2 结构倒格波的波矢与振幅和远程成像的关系	47
3 三维非线性结构设计与全息成像仿真	51
3.1 三维 GS 算法：约束结构角谱为目标图像逆向获得 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 分布	51

3.1.1 基于三维光子晶体的远场傅里叶计算体全息的逆向过程原理.....	51
3.1.2 波长匹配问题.....	52
3.1.3 三维角谱分布函数的初始化.....	54
3.1.4 三维 GS 算法实现过程.....	54
3.2 三维迂回相位编码：对正空间非线性结构多值化编码.....	57
3.2.1 二维迂回相位编码原理.....	58
3.2.2 三维迂回相位编码过程.....	63
3.3 三维光子晶体各面倍频光光强、相位分布及远场成像仿真.....	65
3.3.1 三维光子晶体两侧面的倍频光强分布、相位分布.....	66
3.3.2 三维光子晶体后端面与远场倍频光强分布.....	67
4 总结与展望	70
4.1 总结	70
4.2 展望	71
参考文献	74
致 谢	77
附录 A 中文译文.....	80
附录 B 外文原文.....	88
附录 C 公式推导.....	95
附录 D 线性介质的二阶极化率张量为对称张量.....	159
附录 E 各阶非线性光学系数的性质	162

1 绪论

1.1 引言

1.1.1 非线性光学

1960 年，红宝石激光器神奇的辐射一举扫除了光学领域的沉寂，各种奇异的效应如同雨后春笋展现在世人面前，量子光学、非线性光学、全息技术、超快现象、光学信息处理以及光纤通信等众多崭新的分支学科相继脱颖而出。其中，激光的问世所派生出的傅里叶光学、光学全息技术和非线性光学等新分支尤其令人瞩目。

光与物质相互作用的非线性性质，在 1961 年由 Franken 发现二次谐波以来，逐渐成为如今光学领域最具有活力的学科之一^[1-2]。次年 Bloembergen 等人对非线性光学的参量过程做了详尽的研究，并提出了准相位匹配理论(Quasi phase matching, QPM)，后来被广泛应用于光参量振荡、激光变频等技术中^[3-5]。

1.1.2 非线性光子晶体

非线性光学主要研究光与材料之间的相互作用。随着光学功能的集成化与器件的小型化要求越来越高，天然材料难以满足科技发展需求，于是关于人工微结构材料的研究得到了迅猛发展，比如将材料的折射率进行波长级有序调制，人们制造出了具有光子能带的光子晶体；将金属、聚合物制成亚波长的周期、准周期结构的有序排列，人们发明了超构材料，得到了负折射率或可实现光学隐形等的人工微结构材料。

与上述各人工微结构材料平行地，非线性光子晶体也是一种可实现多种非线性光学效应的人工微结构材料，别名光学超晶格，也称准相位匹配结构，是一种二阶及以上上的非线性系数在空间上受到调制的人工晶体，其最基本的功能是实现倍频或和差频过程中的准相位匹配^[6-8]。通过对结构的特殊设计，非线性光子晶体也能够对非线性过程所产生的倍频光、和差频光的多种属性进行操控，如振幅、相位和轨道角动量等。

非线性光子晶体的调制维度提高后可以在额外的一个维度上弥补非线性过程的波矢失配，进而大大丰富了准相位匹配的类型。准相位匹配的概念最早由诺贝尔奖获得者 Bloembergen 于 1962 年的准相位匹配理论中提出，但直到三十年后光学超晶格的成功制备，它才进入到一个迅速发展的阶段。1998 年，Berger 将准相位匹配结构

从一维推广至二维并正式提出了非线性光子晶体这一称谓，同时也提到了三维非线性光子晶体^[9-10]，但限于当时的技术和制备方法，无法实现这种三维结构^[11]。

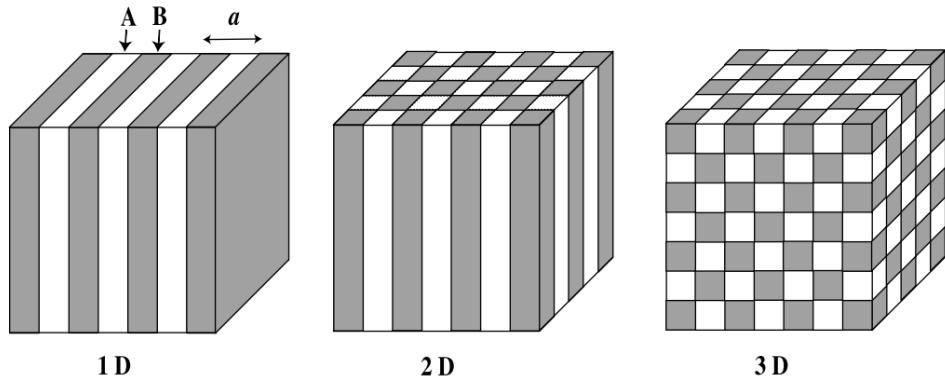


图 1.1 一维、二维、三维非线性光子晶体即受调制的二阶非线性系数空间阵列

2018 年张勇课题组通过飞秒激光直写技术^[12]，通过聚焦光斑破坏局域晶体结构，从而能够在三个空间维度上调制透明铌酸锂晶体的二阶非线性系数，世界首次真正在实验上实现了制备可以实现三维准相位匹配的非线性光子晶体^[13]。三维非线性光子晶体在理论上能真正实现任意角度上的非线性过程的相位匹配，进而实现对经典和量子光波（纠缠光子）的空间模式、频谱、（波包）振幅、相位的全方位调制，展现出在集成量子信息技术、光束的波前调控和波形调控^[14]、非线性全息成像、非线性光信息的储存和重建等领域的广阔应用前景。三维超晶格的制备技术突破也使得光学超晶格的商业化成为可能，为激光波长扩展，全光开关，超快光学等领域奠定了基础。

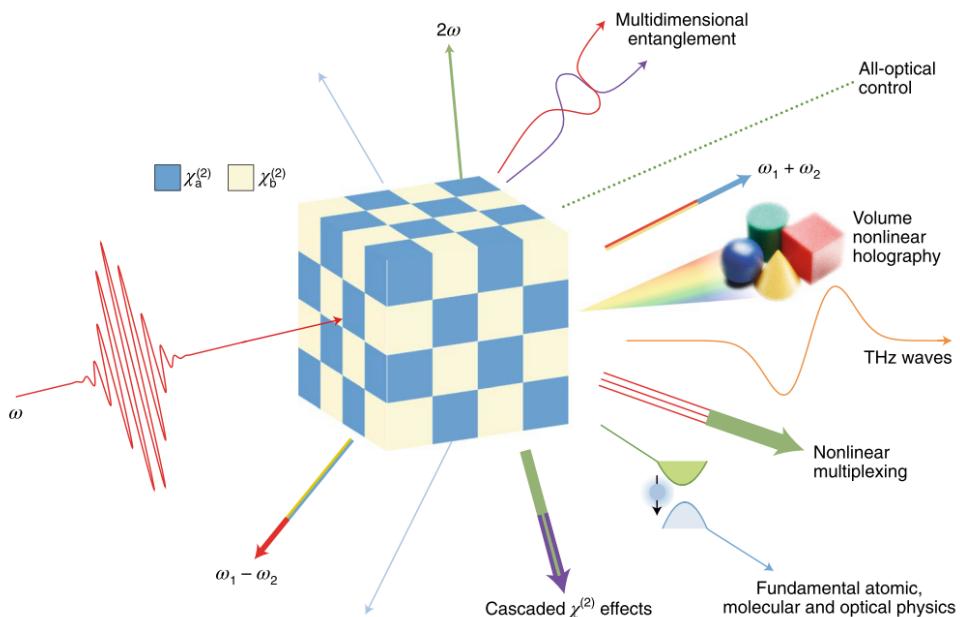


图 1.2 三维非线性光子晶体的潜在应用^[15]

本文基于三维铌酸锂非线性光子晶体，探索其在非线性计算全息成像领域的应用：在前人的准相位匹配理论和非线性全息理论基础上，利用三维 GS 算法设计正空间中三维光子晶体调制区域内的相位分布，并通过三维迂回相位编码法将 GS 算法所得的相位分布转换为晶体实际的二阶非线性系数分布，完成三维非线性光子晶体体全息结构的设计，接着利用数值模拟计算了该三维光子晶体中的倍频过程，同时实现了频率转换和倍频光的三维非线性全息成像。

1.2 非线性全息

“全息”一词来自希腊文“holos”，意为完全的信息。在光学领域是指不仅包括光的振幅信息，还包含光的相位信息。近些年来，随着激光技术的飞速发展，非线性光学与全息技术相结合，取得了许多理论和实验上的突破^[16]。

非线性全息术是一种利用基频光经非线性光子晶体变频后所得的和差倍频光的干涉和衍射原理，记录并再现物体真实的三维图像的技术。非线性全息也延续了普通全息的分类方法，如根据记录介质的厚度可分为平面全息和体全息；根据介质对光波的调制作用，可分为振幅和位相全息；根据物光与参考光在全息图的同侧还是异侧，可分为反射和透射全息；根据全息图记录物光信息的内容不同，可分为近场的菲涅尔全息、远场的傅里叶全息等等。

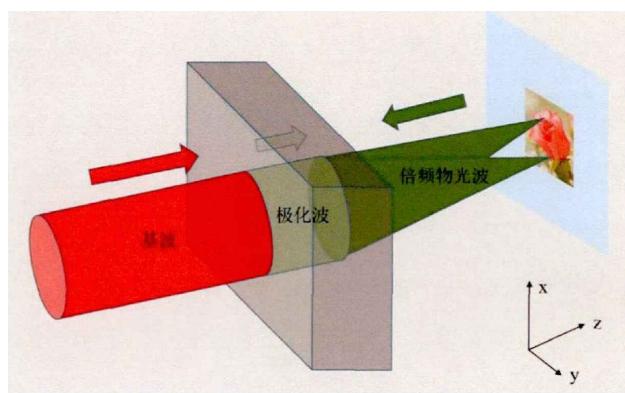


图 1.3 非线性波前记录过程

国内的南京大学继提出了非线性光学中的波动理论——非线性惠更斯原理后，首次提出了非线性菲涅尔全息术，允许通过倍频光重现物光波前：当基波入射非线性晶体后，与介质相互作用时产生非线性倍频极化波，与此同时倍频极化波作为参考光与倍频物光波相互干涉形成干涉条纹，倍频波中所包含的物光信息则被存储在干涉条纹中。由于极化波不能出射晶体，所以没法用晶体外传统的全息板记录该正空间的干涉

条纹；但基波和包含物光信息的重现倍频物光可以出射晶体，并且二者很容易分离。

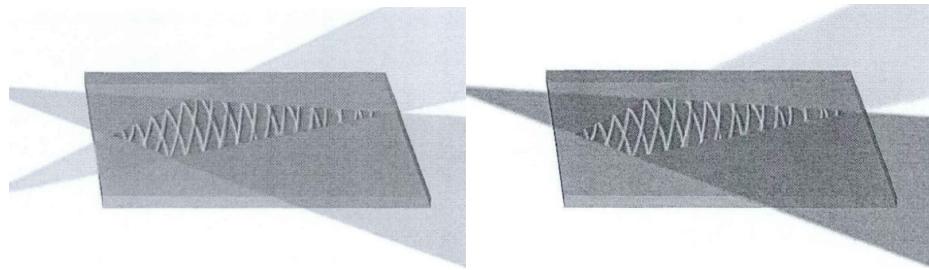


图 1.4 倍频波前记录（左）与再现（右）原理示意图

之后再通过电场极化技术，在非线性晶体中加工出对应干涉条纹的二值化图案，就完成了记录过程。当平面波基频光入射晶体时，其产生的非线性极化波作为参考光，经过晶体内储存的干涉图样畴结构，便可出射与原物光的倍频光的相位和频率一致、振幅成比例的还原倍频物光。此即是非线性全息技术发展早期所提出的非线性菲涅尔全息原理。

1.2.1 非线性全息的三个特征

1.2.1.1 计算全息

不像线性全息，正空间非线性全息的干涉条纹的形成机制是倍频的物光波与倍频的极化波之间的干涉；而由于极化波本身作为参考光不能在真实空间中稳定传播，必须依赖于介质以至时刻处于介质中，因此这种特殊的干涉机制下所形成的干涉条纹，虽然基于正空间的非线性全息原理，是真实存在的，但若不采用特殊手段，该干涉条纹无法被介质以二阶非线性系数分布的形式记录下来，因此由于特殊的干涉机制以及参考光无法在除介质之外的空间中稳定传播，非线性全息均属于计算全息(Computer-generated holography, CGH)的范畴。因此对于非线性全息的全息图，不论是记录物光正空间信息的菲涅尔衍射干涉条纹，还是记录倒空间信息的夫琅禾费衍射干涉条纹，以及三维光子晶体正空间中的目标图像的傅里叶全息图，均只能采用计算机编码的方式记录和生成。而采取计算全息术也有两点优势，一方面不需要在实空间存在具体的对象（但仍需要诸如以二进制形式储存在电脑中的相片等，记录下振幅或复振幅分布的信息），另一方面也可以实现任意理想的波前。

1.2.1.2 体全息

非线性全息除了只能是计算全息之外，一般也尽量采取体全息技术。因为在非线

性光学中，为了有效地提高转换效率，需要增加基波在晶体中的传输距离，于是晶体便不可避免地需要选用三维的人工超晶格；当然，此处的维度只是指提供非线性过程的场所——晶体本身的三维结构是块状还是薄面，并不涉及其内部的二阶非线性系数的调制维度，和可维持波矢匹配、可实现准相位匹配的维度，因此即使是二维光学超晶格（二阶非线性系数是沿某平面二维调制的，在通光方向仍有厚度的“三维”光子晶体），其所参与的二维非线性全息，也属于体全息。

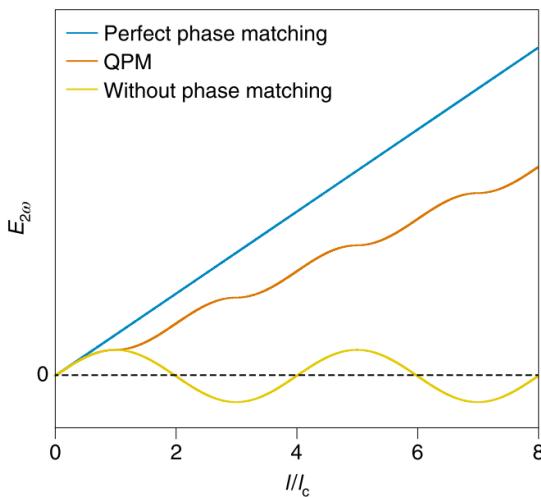


图 1.5 准相位匹配（红线）与完美相位匹配和波矢失配的对比
以及三者的二次谐波光强随传播距离的函数关系

非线性体全息也传承了经典体全息中的选择特性，即因实空间中的非线性惠更斯原理所导出的非线性布拉格衍射，相对于平面全息，体全息的再现光必须严格符合记录时的参考光，包括入射角度、波长等；并且选择特性更强。不同的是线性全息基于光栅的布拉格条件，而非线性全息是基于相位匹配条件。

1.2.1.3 晶体本身同时也是全息板，且需于实验前计算并加工完毕

非线性全息除了必须是计算全息、最好是体全息之外，既可以是菲涅尔全息，又可以是傅里叶全息，但实现准相位匹配的晶体必须同时也是记录全息图像信息的介质——“全息板”。然而一般晶体对于光场干涉条纹的响应主要体现在折射率变化上，而一阶非线性系数的分布无法直接转换为二阶非线性系数分布，这就要求在记录过程需以二阶非线性系数调制的方式将干涉条纹写入非线性光子晶体，而无法通过自然的介质记录干涉图样的方式得到。因此记录工作得分为计算、编码，以及将计算所得的干涉条纹或全息图写入晶体，至少两个步骤。

具体地说，在非线性全息的特殊干涉机制下，正空间中的干涉条纹只在晶体中且只在干涉过程中存在，因此在时空上都很难通过外界的普通的全息板记录下来，以至全息板只能是光子晶体本身。然而任何晶体均无法既吸收光子发生作用以将干涉条纹记录为二阶非线性系数的分布，同时又持续地维持整个准相位匹配过程，因为这是矛盾的——准相位匹配理论告诉我们非线性系数的分布会影响倍频光的生成，而记录过程又要求倍频光的干涉条纹影响介质的二阶非线性系数分布，这样的相互作用总会导致准相位匹配效率或记录效果的其中一方的降低；而干涉条纹在记录为折射率分布的同时却很难从实验上将该分布提取出来并转化为介质的二阶非线性系数分布，种种原因导致非线性全息的记录过程只能是计算全息配合编码加工的形式。

总的来说，在空间上，因特殊的干涉机制导致记录过程中所产生的干涉条纹与光子晶体不可分隔；而为了记录条纹，记录介质与干涉条纹在空间上也不可分离。因此，光子晶体与记录介质在空间上须重叠。但由于晶体无法在时间上做到记录过程与倍频过程的重合，所以记录过程不能采取实验配合化学试剂的方法，而需采用计算配合编码加工的过程，才能将光子晶体加工成以二阶非线性系数记录了全息图信息的介质。

1.2.2 非线性全息的应用

非线性全息一方面在成像领域有着非常重要的地位，另一方面在波前调控的作用也不容小觑^[17-18]。比如利用非线性光学全息的原理，用特殊光束代替物光波，便可以实现任意非线性光束整形。比如将 Airy 光束、拉盖尔-高斯光束、涡旋光束等一系列具有特殊光束与非线性极化波相干涉，就能得到非线性光束整形后的光子晶体结构，进而实现相应非线性光束的输出。

当然，非线性全息还在偏振调控、信息储存以及即时调制器等领域也有非常广阔的应用前景；同时，光学中的非线性全息技术在合适的非线性材料体系中还可以推广到更广泛的电磁波段甚至其它波动体系中去。

1.3 非线性全息编码成像

二维、三维的非线性全息编码成像，都属于计算全息的范畴；但根据“全息板”即非线性光子晶体记录的是倍频物光与极化波参考光的干涉图样，还是以二维物像的傅里叶变换结合编码方法获得的二阶非线性系数的空间分布，仍旧分为菲涅尔计算全息和傅里叶计算全息。

基于非线性菲涅尔全息原理，洪熙昊、陆荣儿等人分别发展了二值化编码法、准多值化编码法（即四值化编码法），将物体经菲涅尔衍射的光场与非线性极化波（参考光）之间的干涉条纹强度分布先二值化或四值化，再利用电场极化翻转铁电材料钽酸锂的相应位置的畴向，以将计算所得的二值化干涉条纹记录在二维非线性晶体中，实验时只需基频重现光照射晶体即可还原倍频物光。但是在这一工作中，由于加工技术限制，只实现了一维光场的复现，这是二维非线性光子晶体的一个缺陷。

具体而言，图 1.1 的左二图所代表的二维光子晶体共有两种成像方式，一种是通光方向沿畴的自发极化方向即 z 向，此时用以补偿色散的结构倒格矢共有 x 向与 y 向分量两种，可以实现 $x-o-y$ 面的准相位匹配和二维非线性全息成像图案的设计，但此时通光方向即 z 向因无法进行结构调制而无法实现准相位匹配，导致无论二维光子晶体 z 向厚度多大，晶体内振幅沿 z 向振荡的倍频光在出射晶体后端面时，振幅只能取下图中无周期结构的 sin 式振荡的最高值，导致倍频效率极低，成像非常暗；第二种方式是选择 x 向或 y 向为通光方向，此时虽可实现通光方向准相位匹配过程提高倍频光场振幅，但由于 z 向准相位匹配过程始终无法实现，导致对 $z-o-x$ 、 $y-o-z$ 面的一维结构调制只能调控一维光场分布，以至于需要采用基频光沿 z 向一层层扫描二维光子晶体，才能实现二维非线性全息成像；并且由于最终生成的二维图像是无数个一维光场的叠加，导致只能做近场菲涅尔全息无法用层状频谱结构做远场傅里叶全息，并且成像图案不可避免会受到相邻层干涉散斑叠加的影响而质量下降。

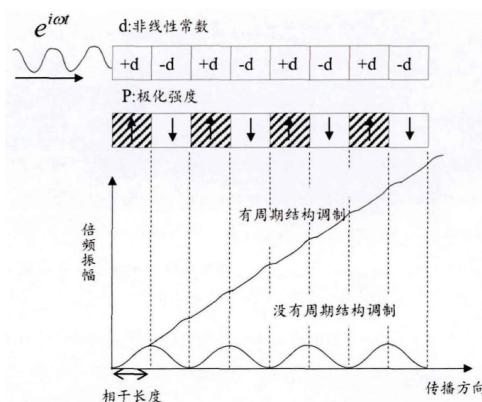


图 1.6 二维非线性光子晶体内周期性反转的铁电畴中准相位匹配过程下的倍频强度分布图

相比而言，基于可提供任意方向的倒格矢弥补波矢失配的三维光子晶体，三维非线性全息没有任何准相位匹配限制，且既可做菲涅尔全息又可实现傅里叶全息。同样，对三维非线性光子晶体内非线性系数分布，也是计算而非实验所得，本文选择傅里叶

全息设计，为了利用三维倒格矢完成准相位匹配同时实现远场成像，采用三维 GS 算法寻找我们所需要倒格矢对应的实空间结构分布，能够使得远场二次谐波光场图案完全满足准相位匹配输出。

1.3.1 二维非线性菲涅尔全息编码成像

在计算得到全息图后，非线性材料将作为全息图的载体，将全息图的内容加工成微结构。在二维光学超晶格材料中，由于其电畴翻转特性，正负畴之间相位相差 π ，因此它可以作为一种天然的二值全息板，同时实现非线性光学效应与全息成像。

人们为了实现二维非线性全息，在基于二维超晶格正负畴结构的基础上，发明了二值化编码方法^[19]，即将干涉光场分布 I_s 取符号函数 $\text{sign}(I_s)$ ，其值随平面光场分布的正负取 +1 或 -1，分别对应待加工超晶格材料中同一平面内的白色正畴和黑色负畴；之后继续发明了准多值化编码方法，但都限于加工工艺的精度限制，以及所选超晶格材料的限制，无法加工出真正连续的畴结构，也就无法应用连续编码；因此二维非线性全息受材料和加工工艺的限制，在二值化→准多值化→多值化→连续编码的路上，遇见了多值化这一瓶颈，成像质量有待进一步提高。同时二维非线性全息的共同特点是无法同时实现二维非线性光场调控的同时实现准相位匹配，这是由于电场极化技术畴翻转必然贯穿整个结构导致的。

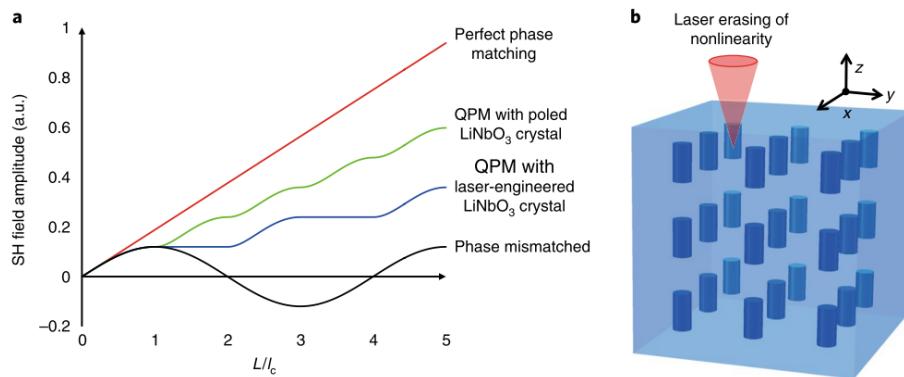


图 1.7 飞秒加工（蓝线）与周期性极化（绿线）的铌酸锂晶体的准相位匹配效率对比^[13]

相应地，虽然三维非线性晶体目前难以实现非线性系数反转（1:-1），但三维非线性结构和三维全息本身相对于二维就多一个维度的自由度，可以提供三个维度上的倒格矢以满足准相位匹配条件，衍生出的相关应用数不胜数；并且由于飞秒直写具有很高的加工精度，即飞秒脉冲对介质的作用时间极短，光斑直径也极小，可以极大地避免在加工过程中产生热效应，防止脉冲能量在介质中扩散，再配合上高精度的移动

装置，使得三维非线性全息在设计和加工上都能近似连续编码；再加上加工过程中也不需要很多其他工艺离不开的掩膜板以及化学制剂，因此可退而求其次只对介质非线性系数的大小进行调制（1:0.85）^[20]，这一思路也曾被运用于多种材料以实现准相位匹配，例如用离子注入的方法降低 GaAs 的非线性系数^[21-24]、用热极化的方法增加玻璃光纤的二阶非线性系数^[25]，以及通过人工设计在超构材料中实现准相位匹配^[26]等。

1.3.2 三维非线性傅里叶全息编码成像

铌酸锂（LiNbO₃）是一种铌、锂和氧的化合物，钛铁矿型结构，铁电相的铌酸锂在光轴方向有个三重对称轴，属于三方晶系，可选择含 6 个 LiNbO₃ 式量的六方晶胞或含 2 个式量的三角晶胞作为晶胞；其在光轴方向还有个三度旋转轴，因此其点群为 3m，空间群为 R3c，不具有对称中心，是压电晶体；同时在其居里温度 1210℃以下，跨过一级相变后随着其温度的降低，电致伸缩能（电弹性能）密度占据总能量密度比例升高，锂离子从原原胞内氧八面体中心相对于氧原子层（ \perp 光轴）沿平行光轴的某个方向位移，同性相斥的缘故导致原顺电相时位于两氧原子层正中且与锂离子共线于光轴的铌离子也随着锂离子相对于氧原子层进行同向位移但位移量不同，导致原胞内正负电中心不重合，以至于出现宏观和微观的自发极化，并且单畴态的铌酸锂的自发极化方向与其光轴方向平行或反平行，随着温度的降低自发极化程度提高，有热电效应，此时铌酸锂具有自发极化且电畴方向能随外加电场方向的改变转向而是铁电晶体，且易于加工；同时氧八面体中心存在锂离子空位而属于 p 型半导体；而从一阶非线性系数上，它还是负单轴晶体（ $n_o > n_e$ ）。经过畸化处理的铌酸锂晶体具有压电、铁电、热电、光电、非线性光学等多性能。其单晶是光波导、移动电话、压电传感器、光学调制器和各种其它线性光学应用的重要材料，同时因其非线性系数 d_{33} 和光损伤阈值均较高，所以也是非线性光学应用的重要材料。

本文的研究所采用的技术是飞秒激光直写技术以调制铌酸锂晶体的二阶非线性系数的三维分布，而非采用光刻的方法镀上周期性电极和二倍周期地反向施加脉冲电压^[27-28]以制成周期极化畴反转的二维准相位匹配结构，因此调制出来的是三维光子晶体；正是因为由于采用的加工工艺、非线性全息的类别、人造微结构材料及其调制维度的不同，采用的数值计算的对象和方法、计算全息编码方法也不一样，相应地我们采用三维 GS 算法先算得正空间相位三维分布即傅里叶相位体全息图，再将三维的相位分布通过三维迂回相位编码法转换为二阶非线性系数的三维分布，并且由于飞秒激

光的加工精度允许，对相位的编码是多值化且准连续的。最后用基频光入射晶体并采用二级倍频光成像。

相比而言，这种加工技术及编码方法在效率上比上述两种二维非线性全息高，且因准相位匹配维度的提高而具有更高的信息容量，即三维非线性全息可产生二维图像，而二维非线性全息只可生成一维光场分布；在普通的二维非线性全息中存在畴极化方向的相位失配导致非线性转换效率较低，而基于三维非线性光子晶体的非线性体全息在该方向的转换效率上具有天然的优势，毕竟非线性光子晶体及非线性全息从二维发展到三维是质的创新。三维非线性全息也能在高效频率转换的同时对光束整形和重现全息图像，即全方位调控光场，这也是线性全息和二维非线性全息所不可比拟的。

2 三维非线性傅里叶全息成像原理

2.1 非线性光学

激光的问世使作用于介质的光场振幅提高了好几个数量级，大量涌现出的旧理论无法解释的新现象打破了介质电极化强度与介质中外电场一次方呈简单线性关系的旧有认识：

$$\begin{aligned} \tilde{\mathbf{P}}_n(\mathbf{r}, t) = \tilde{\mathbf{P}}(\omega_n) &= \varepsilon_0 \left[\sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}}} \chi_e^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}(\omega_{m_{\pi(1)}}) + \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}}} D_{m_1, m_2} \chi_e^{(2)}(\omega_n; \omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}) \right. \\ &\quad : \tilde{\mathbf{E}}(\omega_{m_{\pi(1)}}) \tilde{\mathbf{E}}(\omega_{m_{\pi(2)}}) + \dots + \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}} + \dots + \omega_{m_{\pi(l)}}} D_{m_1, m_2, \dots, m_l} \overbrace{\chi_e^{(l)}(\omega_n; \overbrace{\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}, \dots, \omega_{m_{\pi(l)}}}^l)}^{\overbrace{\tilde{\mathbf{E}}(\omega_{m_{\pi(1)}}) \tilde{\mathbf{E}}(\omega_{m_{\pi(2)}}) \cdots \tilde{\mathbf{E}}(\omega_{m_{\pi(l)}})}^l} + \dots] \end{aligned} \quad (2.1)$$

其中，求和区间遍历符合条件 $\omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}} = \omega_n$ 的单色光场的排列 $(\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}})$ 的集合 $\{(\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}})\}$ （以 $l=2$ 为例），下标 n, m 对应的 $\{\omega_n\}, \{\omega_m\}$ 分别表示电极化强度 $\tilde{\mathbf{P}}$ 和电场强度 $\tilde{\mathbf{E}}$ 各自的频谱；并规定对 $\tilde{\mathbf{E}}$ 的频谱 $\{\omega_m\}$ 而言， $\{m\} \leftrightarrow \{\omega_m\}$ ；同理对 $\tilde{\mathbf{E}}$ 的单色光波集 $\{\omega_m\}$ ，也有 $\{i\} \leftrightarrow \{m_i\} \leftrightarrow \{\omega_{m_i}\}$ ；且二者之间满足： $\{i\} \leftrightarrow \{m_i\} \xrightarrow[\text{但满射}]{\text{非单射}} \{m_i\} \leftrightarrow \{m\}$ ，即即使 $\{i\}$ 中 $i_1 \neq i_2$ ，也可以有 $m_{i_1} = m_{i_2}$ ，此时 m_{i_1}, m_{i_2} 在 $\{m_i\}$ 中值相同，简并为同一频率；在 $\{m_i\}$ 中值不同，表频率相同的两束单色光，可见 $\{i\}$ 与 $\{m\}$ 不一一映射。说明 $\tilde{\mathbf{E}}$ 的{单色光} $\xrightarrow[\text{满射}(\checkmark)]{\text{单射}(x)} \tilde{\mathbf{E}}$ 的{频率}，即频率相同的多束参与作用的 $\tilde{\mathbf{E}}_i$ ，在 $\{\omega_{m_i}\}$ 内仍是互不同的元素 ω_{m_i} ，但在 $\{\omega_m\}$ 内简并为同一 ω_m 。

求和项中包含了 $\omega_{m_{\pi(1)}} = \omega_{m_{\pi(2)}}$ 的情况，此时相应项即描述(二)倍频极化光，并在相位匹配条件满足以及结构的倒格矢可补上相应波矢失配时，作为非线性介质中的波动方程等式右侧的非线性波源项，将在介质中对应地持续产生二次谐波，否则

$\omega_{m_{\pi(1)}} \neq \omega_{m_{\pi(2)}}$ 即描述的是产生和频光或差频光。

若当 $m_1 \neq m_2$ 时， $(\omega_{m_{\pi(2)}}, \omega_{m_{\pi(1)}}) \neq (\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}})$ ，即集合 $\{(\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}})\}$ 中的每个元素是一种“排列性序列”，则 $D_{m_1, m_2} = 1$ ；否则若任一排列光场序列 $(\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}, \dots, \omega_{m_{\pi(l)}})$ 对于任意两个独立取值且值不相同的角标 m_{i_1}, m_{i_2} 均具有置换对称性，即若满足 $\omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}} + \dots + \omega_{m_{\pi(l)}} = \omega_n$ 的集合 $\{(\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}, \dots, \omega_{m_{\pi(l)}})\}$ 中每个元素是一种“组合性序列”，则：

$$D_{m_1, m_2, \dots, m_l} = \frac{l!}{\prod_k^{\{m\}} \text{card}\{i_{m=k}\}!} = \frac{l!}{\prod_{k=1}^{m_{\max}} \text{card}\{i_{m=k}\}!} = \frac{l!}{\sum_k^{\{m\}} \text{card}\{\overline{m_i | m_i = k}\}!} = \frac{l!}{\sum_m \text{card}\{\overline{m_i | m_i = m}\}!} \quad (2.2)$$

它称为简并因子，体现了倍频光与和差频光的2倍振幅区别，也体现了参与转换的是光而非频率。

其中 D_{m_1, m_2, \dots, m_l} 内及各处数学符号的来源，需参考附 C1.1.1，下图也可帮助理解各数学符号的物理意义：

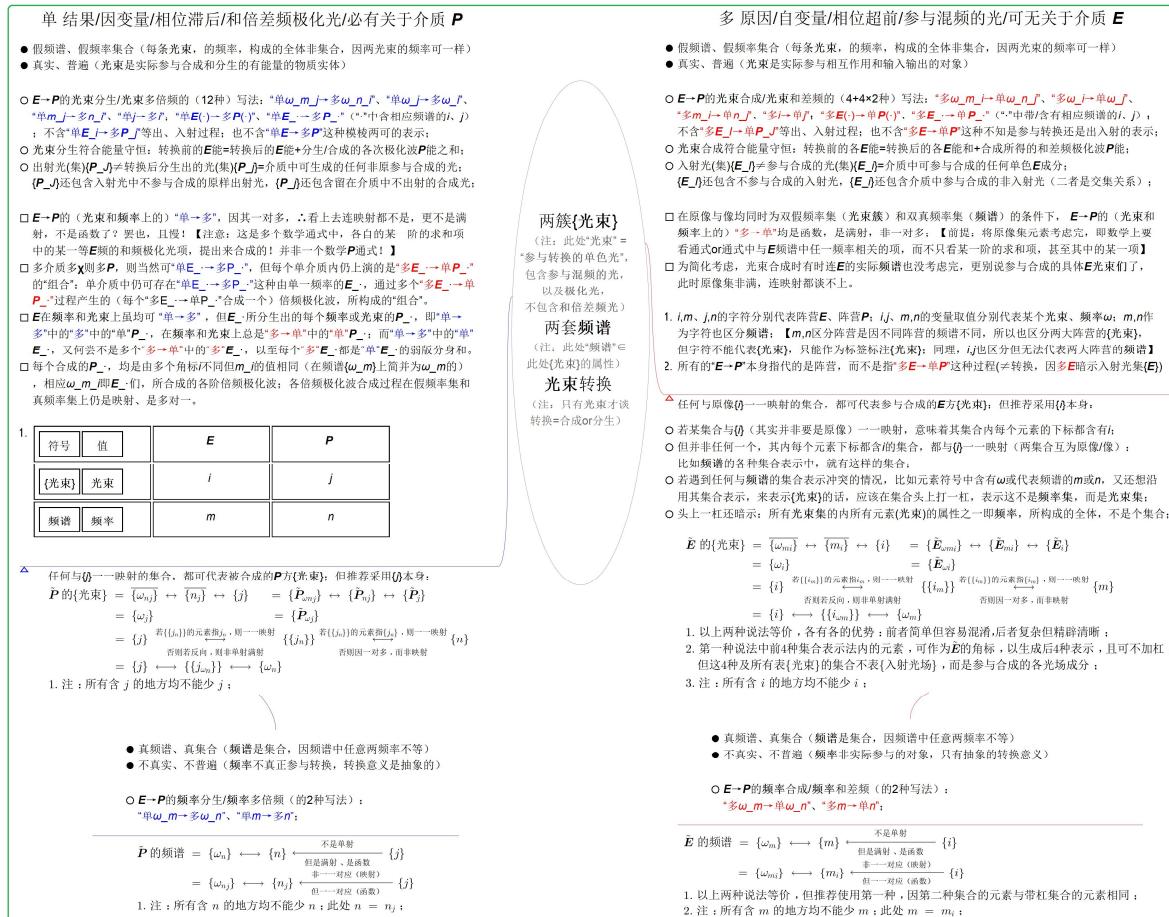


图 2.1 非线性过程思维导图图解

更常用地，在推导一维条件下非线性电介质中的波动方程、耦合波方程和相应的相位匹配问题和准相位匹配条件时，采用的是分量形式或标量形式下的电极化强度表达式，需注意其中 $\tilde{P}_{u_1}(\omega_n) = \tilde{P}_{nu_1}(\mathbf{r}, t) \neq \tilde{P}_n(u_1, t)$ ：

$$\tilde{P}_{u_1}(\omega_n) = \epsilon_0 \left[\sum_{\pi}^{\omega_n = \omega_{m_{\pi}(1)}} \sum_{u_2}^{\{i,j,k\}} \chi_{u_1 u_2}^{(1)}(\omega_n) \tilde{E}_{u_2}(\omega_{m_{\pi}(1)}) + \sum_{\pi}^{\omega_n = \omega_{m_{\pi}(1)} + \omega_{m_{\pi}(2)} + \dots + \omega_{m_{\pi}(l)}} D_{m_1, m_2, \dots, m_l} \right]$$

$$\cdot \sum_{u_2 u_3}^{\{i,j,k\}} \chi_{u_1 u_2 u_3}^{(2)}(\omega_n; \omega_{m_{\pi}(1)}, \omega_{m_{\pi}(2)}) \tilde{E}_{u_2}(\omega_{m_{\pi}(1)}) \tilde{E}_{u_3}(\omega_{m_{\pi}(2)}) + \dots + \sum_{\pi}^{\omega_n = \omega_{m_{\pi}(1)} + \omega_{m_{\pi}(2)} + \dots + \omega_{m_{\pi}(l)}} D_{m_1, m_2, \dots, m_l}$$

$$\cdot \sum_{u_2 u_3 \cdots u_{l+1}}^{\{i,j,k\}} \chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \overbrace{\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}, \dots, \omega_{m_{\pi(l)}}}^l) \overbrace{\tilde{E}_{u_2}(\omega_{m_{\pi(1)}}) \tilde{E}_{u_3}(\omega_{m_{\pi(2)}}) \cdots \tilde{E}_{u_{l+1}}(\omega_{m_{\pi(l)}})}^l + \dots] \quad (2.3)$$

其中 $\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}$ 有 $l+1$ 个协变指标，且每个指标分别可取 i, j, k 三值，因此它确是个 3 维 $l+1$ 阶张量，有 3^{l+1} 个分量；比如在最低阶的二阶非线性效应 $\tilde{\mathbf{P}}^{(2)}(\omega_n)$ 中， $\chi_{u_1 u_2 u_3}^{(2)}$ 仍有足足 27 个分量，尽管可以通过选择特定的坐标系、选择具有一定对称性的晶体、满足特殊的入射情况，或者从无色散无损耗介质中的克莱门对称性的角度，等手段使得其中一些分量归零，以减少二阶非线性系数矩阵中的独立分量个数（见附录 E）。

2.1.1 从麦氏方程组的微分形式到非线性电介质中的波动方程

可以证明，普遍形式(介质中)的麦氏方程组的微分形式，(其内四个方程均)仍适用于非均匀、各向异性，甚至非线性的电磁介质，并且适用于非稳恒(电磁)场。以 SI 单位制写为：

$$\left\{ \begin{array}{l} \nabla \cdot \tilde{\mathbf{D}} = \tilde{\rho}_f \\ \nabla \cdot \tilde{\mathbf{B}} = 0 \\ \nabla \times \tilde{\mathbf{E}} = -\frac{\partial \tilde{\mathbf{B}}}{\partial t} \\ \nabla \times \tilde{\mathbf{H}} = \tilde{\mathbf{J}}_f + \frac{\partial \tilde{\mathbf{D}}}{\partial t} \end{array} \right. \quad (2.4)$$

其中 $\tilde{\mathbf{D}}, \tilde{\mathbf{B}}, \tilde{\mathbf{E}}, \tilde{\mathbf{H}}$ 都是复色场、时变场。在一般的电介质中，自由电荷体密度 $\tilde{\rho}_f = 0$ ，因此 $\tilde{\mathbf{J}}_f = \tilde{\rho}_f \cdot \mathbf{v} = \mathbf{0}$ ，则四号方程右侧只剩极化电流与位移电流之和 $\partial \tilde{\mathbf{D}} / \partial t$ ；若所考虑的材料是磁性的，则 $\tilde{\mathbf{B}} = \mu_0 (\tilde{\mathbf{H}} + \tilde{\mathbf{M}})$ 。将考虑上述两个条件的四号方程，代入经 $\nabla \times$ 作用后的三号方程中，得：

$$\begin{aligned} \nabla \times (\nabla \times \tilde{\mathbf{E}}) &= \nabla \times \left(-\frac{\partial \tilde{\mathbf{B}}}{\partial t} \right) = -\frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{B}}) \xrightarrow{\text{四号方程}} -\mu_0 \frac{\partial}{\partial t} \left(\frac{\partial \tilde{\mathbf{D}}}{\partial t} + \nabla \times \tilde{\mathbf{M}} \right) \\ &= -\mu_0 \left[\frac{\partial^2 \tilde{\mathbf{D}}}{\partial t^2} + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}) \right] \xrightarrow[c=\sqrt{\frac{1}{\mu_0 \epsilon_0}}]{\tilde{\mathbf{D}} = \epsilon_0 \tilde{\mathbf{E}} + \tilde{\mathbf{P}}} - \frac{1}{c^2 \cdot \mu_0 \epsilon_0} \cdot \mu_0 \left[\frac{\partial^2 (\epsilon_0 \tilde{\mathbf{E}} + \tilde{\mathbf{P}})}{\partial t^2} + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}) \right] \\ &= -\frac{1}{c^2} \frac{\partial^2 \tilde{\mathbf{E}}}{\partial t^2} - \frac{1}{c^2 \cdot \epsilon_0} \left[\frac{\partial^2 \tilde{\mathbf{P}}}{\partial t^2} + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}) \right] \end{aligned} \quad (2.5)$$

得非线性介质中电场波动方程的最普遍形式：

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}) + \frac{1}{c^2} \frac{\partial^2 \tilde{\mathbf{E}}}{\partial t^2} = -\frac{1}{\epsilon_0 c^2} \left[\frac{\partial^2 \tilde{\mathbf{P}}}{\partial t^2} + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}) \right] \quad (2.6)$$

2.1.1.1 复数形式傅里叶展开以单色化叠加场、分离时间项、引入电磁非线性效应

但上述各物理量都是叠加场，为引入非线性效应以及分离时间项以将各量和波动方程变为定态的，首先需要将方程和物理量全都单色化。现选择合适的时空坐标系、时空坐标原点，以及延拓方法，使得延拓后的 $\tilde{\mathbf{E}}(t)$ 、 $\tilde{\mathbf{E}}(\mathbf{r})$ 以及 $\tilde{\mathbf{E}}(\mathbf{r}, t)$ 在四个维度上分别都是偶函数且是周期函数，则延拓后的叠加场可展为傅里叶余弦级数，接着再将其转换为复数形式傅里叶级数展开式，将参与混频的复色场展为单色场的线性叠加：

(1) 参与混频的复色场 $\tilde{\mathbf{E}}(\mathbf{r}, t)$ 的单色化，并分离其时间项

$$\begin{aligned} \tilde{\mathbf{E}}(\mathbf{r}, t) &= \sum_{i=0}^{\infty} \mathcal{E}_i \cos(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t) = \sum_{i=-\infty}^{\infty} A_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} = \sum_{i=-\infty}^{\infty} \tilde{\mathbf{E}}_i(\mathbf{r}, t) \\ &\left(\mathbf{A}_0 = \mathcal{E}_0, \mathbf{A}_i = \frac{\mathcal{E}_{|i|}}{2} (i \neq 0); |A| \in \mathbb{R} \right) \end{aligned} \quad (2.7)$$

其中定义取值为负整数的 i 及其所代表的波的各项光学参数 ω_i, \mathbf{k}_i 和 A_i, \mathbf{E}_i 连同 $\tilde{\mathbf{E}}_i$ 本身的物理意义：对于 $i > 0$ ，定义 $\omega_{-i} := -\omega_i$ 、 $\mathbf{k}_{-i} := -\mathbf{k}_i$ ，于是：

$$\tilde{\mathbf{E}}_i^* = \tilde{\mathbf{E}}(\omega_i, \mathbf{k}_i)^* = \tilde{\mathbf{E}}_i(\mathbf{r}, t)^* = \mathbf{E}_i(\mathbf{r})^* e^{-i\omega_{-i} t} \quad (2.8)$$

接着再定义：

$$A_i^* = \mathbf{A}(\omega_i, \mathbf{k}_i)^* := \mathbf{A}(-\omega_i, -\mathbf{k}_i) = \mathbf{A}(\omega_{-i}, \mathbf{k}_{-i}) = A_{-i} \quad (2.9)$$

则有：

$$\mathbf{E}_i^* = \mathbf{E}(\omega_i, \mathbf{k}_i)^* = \mathbf{A}(\omega_i, \mathbf{k}_i)^* e^{i(-\mathbf{k}_i) \cdot \mathbf{r}} = \mathbf{A}(\omega_{-i}, \mathbf{k}_{-i}) e^{i\mathbf{k}_{-i} \cdot \mathbf{r}} = \mathbf{E}(\omega_{-i}, \mathbf{k}_{-i}) = \mathbf{E}_{-i} \quad (2.10)$$

反代回 $\tilde{\mathbf{E}}_i^*$ 得：

$$\tilde{\mathbf{E}}_i^* = \tilde{\mathbf{E}}(\omega_i, \mathbf{k}_i)^* = \mathbf{E}(\omega_{-i}, \mathbf{k}_{-i}) e^{-i\omega_{-i} t} = \tilde{\mathbf{E}}(\omega_{-i}, \mathbf{k}_{-i}) = \tilde{\mathbf{E}}_{-i} \quad (2.11)$$

这样定义的目的在于原具有物理意义的实数形式单色子波本只能写为：

$$\tilde{\mathcal{E}}_i = \mathcal{E}_i \cos(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t) = A_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} + A_i e^{-i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} \quad (2.12)$$

但现在可以直接写作：

$$\tilde{\mathcal{E}}_i = \tilde{\mathbf{E}}_i + \tilde{\mathbf{E}}_{-i} = \tilde{\mathbf{E}}_i + \tilde{\mathbf{E}}_i^* \quad (2.13)$$

为之后单色能流密度等可观测量的写法加以简化埋下伏笔。

(2) 单色化非线性、线性过程所产生的复色极化场 $\tilde{\mathbf{P}}(\mathbf{r}, t)$ ，并分离其时间项

现考虑“多 $\tilde{\mathbf{E}}_i \rightarrow$ 单 $\tilde{\mathbf{P}}_n$ ”的非线性光学过程，考虑某一频率为 ω_n 的单色和频极化光，它作为目标即混频/合成的结果，本身就是由式 1-1 中各阶非线性效应 $\tilde{\mathbf{P}}_n^{(l)}$ 叠加而

成，即 $\tilde{\mathbf{P}}_n(\mathbf{r}, t) = \sum_{l=1}^{\infty} \tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t)$ ，而对于其中每一个 $\tilde{\mathbf{P}}_n^{(l)}$ 成分，都有 l 束满足能量守恒 $\sum_{i=1}^l \omega_i = \omega_n$ 的单色 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 参与合成，因此每个合成的极化光 $\tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t) = \mathbf{P}_n^{(l)}(\mathbf{r}) e^{-i\omega_n t}$ 的时间部分为 $e^{-i\omega_n t} = e^{-i\sum_{i=1}^l \omega_i t} = \prod_{i=1}^l e^{-i\omega_i t}$ ，其连同频率 ω_n 均不随阶数 l 的变化而变化，是个固定值。因此合成的 $\tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t)$ 再合成所得的叠加场（虽是叠加场，但仍是单色的） $\tilde{\mathbf{P}}_n(\mathbf{r}, t) = \sum_{l=1}^{\infty} \tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t) = e^{-i\omega_n t} \sum_{l=1}^{\infty} \mathbf{P}_n^{(l)}(\mathbf{r}) = \mathbf{P}_n(\mathbf{r}) e^{-i\omega_n t}$ 的时间部分，是所有单色 $\tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t)$ 的公因式 $e^{-i\omega_n t}$ ，也是随所考虑的频率即 n 的选定而固定了的。

接着考虑多个“多 $\tilde{\mathbf{E}}_i \rightarrow$ 单 $\tilde{\mathbf{P}}_n$ ”过程，合成多个不同频率 ω_n （等价于不同光束 ω_{n_j} ）的和/差频极化光 $\tilde{\mathbf{P}}_n(\mathbf{r}, t)$ ，它所叠加起来的复合场便与式(2.7)中的 $\tilde{\mathbf{E}} = \sum_{i=-\infty}^{\infty} \tilde{\mathbf{E}}_i$ 拥有同等地位：

$$\begin{aligned} \tilde{\mathbf{P}}(\mathbf{r}, t) &= \sum_{n=-\infty}^{\infty} \tilde{\mathbf{P}}_n(\mathbf{r}, t) = \sum_{n=-\infty}^{\infty} \sum_{l=1}^{\infty} \tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t) \\ &= \sum_{n=-\infty}^{\infty} e^{-i\sum_{i=1}^l \omega_i t} \sum_{l=1}^{\infty} \mathbf{P}_n^{(l)}(\mathbf{r}) = \sum_{n=-\infty}^{\infty} \mathbf{P}_n(\mathbf{r}) e^{-i\omega_n t} \end{aligned} \quad (2.14)$$

(3) 单色化非线性、线性过程所产生的复色电位移矢量场 $\tilde{\mathbf{D}}(\mathbf{r}, t)$ ，分离时间项据各阶单色非线性电位移矢量与相应阶的非线性单色极化波的关系：

$$\tilde{\mathbf{D}}_n = \varepsilon_0 \overset{\rightarrow}{\epsilon}_r^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}_n = \overset{\rightarrow}{\epsilon}^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}_n \quad (2.15)$$

$$\tilde{\mathbf{D}}_n^{(l)} = \tilde{\mathbf{P}}_n^{(l)} (l \geq 2) \quad \tilde{\mathbf{D}}_n^{NL} = \tilde{\mathbf{P}}_n^{NL} = \sum_{l=2}^{\infty} \tilde{\mathbf{P}}_n^{(l)} \quad \tilde{\mathbf{D}}_n = \tilde{\mathbf{D}}_n^{(1)} + \tilde{\mathbf{D}}_n^{NL} \quad \tilde{\mathbf{P}}_n = \tilde{\mathbf{P}}_n^{(1)} + \tilde{\mathbf{P}}_n^{NL} \quad (2.16)$$

有：

$$\begin{aligned} \tilde{\mathbf{D}}(\mathbf{r}, t) &= \sum_{n=-\infty}^{\infty} \tilde{\mathbf{D}}_n(\mathbf{r}, t) = \sum_{n=-\infty}^{\infty} \sum_{l=1}^{\infty} \tilde{\mathbf{D}}_n^{(l)}(\mathbf{r}, t) \\ &= \sum_{n=-\infty}^{\infty} e^{-i\sum_{i=1}^l \omega_i t} \sum_{l=1}^{\infty} \mathbf{D}_n^{(l)}(\mathbf{r}) = \sum_{n=-\infty}^{\infty} \mathbf{D}_n(\mathbf{r}) e^{-i\omega_n t} \end{aligned} \quad (2.17)$$

在求和 $\tilde{\mathbf{P}} = \sum_{n=-\infty}^{\infty} \tilde{\mathbf{P}}_n$ 、 $\tilde{\mathbf{D}} = \sum_{n=-\infty}^{\infty} \tilde{\mathbf{D}}_n$ 中，也定义了负角标所对应的负频率、负波矢以及振幅部分、复振幅部分和相应时变物理量的共轭量 $\tilde{\mathbf{P}}_n^* = \tilde{\mathbf{P}}_{-n}$ 、 $\tilde{\mathbf{D}}_n^* = \tilde{\mathbf{D}}_{-n}$ ，以方便实场 $\tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 的描述。现考虑如下情景：由于负频的原料光集 $\{\tilde{\mathbf{E}}_i^*\}$ 理应也通过式(2.1)，以各个负频子波的加和的方式，合成出负频的和频极化光 $\tilde{\mathbf{P}}_n^*$ ；同时将正频 $\{\tilde{\mathbf{E}}_i\}$ 到正频 $\tilde{\mathbf{P}}_n$ 的式(2.1)两旁取复共轭，并将结果与前者比较，得出 $\overset{\rightarrow}{\chi}_e^{(l)}(\omega_n; \omega_{m_1}, \omega_{m_2}, \dots, \omega_{m_l}) = \overset{\rightarrow}{\chi}_e^{(l)}(-\omega_n; -\omega_{m_1}, -\omega_{m_2}, \dots, -\omega_{m_l})$ ，即同阶非线性系数的复共轭与其关于除负频外相同入射光集的函数值相等。有关各阶非线性系数的性质可查阅附录 E。

既然 $\tilde{\mathbf{P}}_n^{(l)}$ 与 $\tilde{\mathbf{D}}_n^{(l)} = \tilde{\mathbf{P}}_n^{(l)} (l \geq 2)$ 中的时间部分是通过式(2.1)以 $e^{-i\omega_n t} = e^{-i\sum_{i=1}^l \omega_i t} = \prod_{i=1}^l e^{-i\omega_i t}$ 的方式产生的，而其空间部分中的相位部分，即波矢部分，由于与时间部分即频率部

分的数学形式相同、地位相同，那么在混频过程中，若同一束 l 阶单色极化波的频率部分由 $e^{-i\omega_n t} = e^{-i\sum_{i=1}^l \omega_i t} = \prod_{i=1}^l e^{-i\omega_i t}$ 给定，则其波矢部分也将由光波序列同为 $\{i\}$ 的相同的 l 束单色光的波矢部分相乘而得： $e^{-ik_n \cdot r} = e^{-i\sum_{i=1}^l k_i \cdot r} = \prod_{i=1}^l e^{-ik_i \cdot r}$ ，因此在 $\tilde{\mathbf{P}}_n^{(l)}$ 与 $\tilde{\mathbf{D}}_n^{(l)}$ 的生成过程中，这 $l+1$ 束波除了满足能量/频率守恒 $\sum_{i=1}^l \omega_i = \omega_n$ 外，还满足动量/波矢守恒 $\sum_{i=1}^l \mathbf{k}_i = \mathbf{k}_n$ ，也就是说 $\tilde{\mathbf{P}}_n^{(l)}$ 与 $\tilde{\mathbf{D}}_n^{(l)}$ 的复振幅部分 $\mathbf{P}_n^{(l)}(\mathbf{r}), \mathbf{D}_n^{(l)}(\mathbf{r})$ 中必然含有相位因子 $e^{-ik_n \cdot r}$ ，且其中波矢 $\mathbf{k}_n = \sum_{i=1}^l \mathbf{k}_i$ ；同理， $\tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 及 $\mathbf{P}_n(\mathbf{r}), \mathbf{D}_n(\mathbf{r})$ 其中也应均含有 $e^{-ik_n \cdot r}$ 且 $\mathbf{k}_n = \sum_{i=1}^l \mathbf{k}_i$ ，但由于其中包含 $\tilde{\mathbf{P}}_n^{(1)}, \tilde{\mathbf{D}}_n^{(1)}$ ，而二者中又含有 $\tilde{\mathbf{E}}_n$ ，所以对于 $\tilde{\mathbf{E}}_n, \tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ ，事情要稍微复杂一点。

(4) 各阶非线性极化波源所产生的同频非线性和/差频波 $\tilde{\mathbf{E}}_n(\mathbf{r}, t)$ 的时空部分

当 $\tilde{\mathbf{E}}_n$ 中的时间部分为 $e^{-i\omega_n t} = e^{-i\sum_{i=1}^l \omega_i t}$ ，且空间部分为 $e^{-ik_n \cdot r} = e^{-i\sum_{i=1}^l k_i \cdot r}$ 时，称为(符合)完美相位匹配条件，但一般情况下只可能符合其中一个条件；为将波动方程化为定态以便求解，一般先假设非线性波源所生成的和频波的频率满足 $\omega_n = \sum_{i=1}^l \omega_i$ ，并于波动方程化为定态且解出 $\tilde{\mathbf{E}}_n$ 后，再考察波矢匹配与否以及相应的结果，即是否有 $\tilde{\mathbf{E}}_n$ 中的 $e^{-ik_n \cdot r}$ 中的 $\mathbf{k}_n = \sum_{i=1}^l \mathbf{k}_i$ ，此时其 \mathbf{k}_n 却不一定 $= \sum_{i=1}^l \mathbf{k}_i$ 。之后会给出，这是一个条件，而不是一个结论或事实；相反，要同时实现波矢和频率匹配，即完美相位匹配，是极其苛刻的。

对于为将方程变为定态而将方程的解的频率设为与 $\tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 的频率 ω_n 相同的 $\tilde{\mathbf{E}}_n = \mathbf{E}_n(\mathbf{r})e^{-i\omega_n t}$ ，也定义有 $\tilde{\mathbf{E}}_n^* = \tilde{\mathbf{E}}_{-n}$ ，只不过要注意 $\tilde{\mathbf{E}}_n$ 中的 \mathbf{k}_n 与 $\tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 中的 \mathbf{k}_n 含义不同，且值可能相同或不同。

(5) 磁场波动方程、单色化 $\tilde{\mathbf{H}}, \tilde{\mathbf{M}}, \tilde{\mathbf{B}}$ 、引入 $\tilde{\mathbf{M}}, \tilde{\mathbf{B}}$ 的非线性效应

类似电场波动方程(2.6)地， $\nabla \times$ 无源条件下的四号方程可得非线性介质中磁场波动方程的最普遍形式：

$$\nabla \times (\nabla \times \tilde{\mathbf{H}}) + \frac{1}{c^2} \frac{\partial^2 \tilde{\mathbf{H}}}{\partial t^2} = -\frac{1}{c^2} \frac{\partial^2 \tilde{\mathbf{M}}}{\partial t^2} + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{P}}) \quad (2.18)$$

对于式(2.6)式(2.18)中的复色场 $\tilde{\mathbf{M}}$ ，连同 $\tilde{\mathbf{B}}$ 均可分解为线性部分和非线性部分：

$$\begin{aligned} \tilde{\mathbf{B}}_n^{(1)} &= \mu_0 [\tilde{\mathbf{H}}_n + \tilde{\mathbf{M}}_n^{(1)}] = \mu_0 [\tilde{\mathbf{H}}_n + \overleftrightarrow{\chi}_m^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}_n] = \mu_0 [\overleftrightarrow{\mathbf{I}} + \overleftrightarrow{\chi}_m^{(1)}(\omega_n)] \cdot \tilde{\mathbf{H}}_n \\ &= \mu_0 \overrightarrow{\boldsymbol{\mu}}_r^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}_n = \overrightarrow{\boldsymbol{\mu}}^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}_n \end{aligned} \quad (2.19)$$

$$\tilde{\mathbf{B}}_n^{(l)} = \mu_0 \tilde{\mathbf{M}}_n^{(l)} \quad (l \geq 2) \quad , \quad \tilde{\mathbf{B}}_n^{NL} = \mu_0 \tilde{\mathbf{M}}_n^{NL} = \mu_0 \sum_{l=2}^{\infty} \tilde{\mathbf{M}}_n^{(l)}$$

$$\tilde{\mathbf{B}}_n = \tilde{\mathbf{B}}_n^{(1)} + \tilde{\mathbf{B}}_n^{NL} \quad , \quad \tilde{\mathbf{M}}_n = \tilde{\mathbf{M}}_n^{(1)} + \tilde{\mathbf{M}}_n^{NL} \quad (2.20)$$

其中类似式(2.1)地，非线性磁效应可以表示为：

$$\begin{aligned} \tilde{\mathbf{M}}_n(\mathbf{r}, t) = \tilde{\mathbf{M}}(\omega_n) &= \sum_{\pi}^{\omega_n=\omega_{m_{\pi(1)}}} \overrightarrow{\chi}_m^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}(\omega_{m_{\pi(1)}}) + \sum_{\pi}^{\omega_n=\omega_{m_{\pi(1)}}+\omega_{m_{\pi(2)}}} D_{m_1, m_2} \overrightarrow{\chi}_m^{(2)}(\omega_n; \omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}) \\ &\quad : \tilde{\mathbf{H}}(\omega_{m_{\pi(1)}}) \tilde{\mathbf{H}}(\omega_{m_{\pi(2)}}) + \dots + \sum_{\pi}^{\omega_n=\omega_{m_{\pi(1)}}+\omega_{m_{\pi(2)}}+\dots+\omega_{m_{\pi(l)}}} D_{m_1, m_2, \dots, m_l} \overbrace{\chi_m^{(l)}(\omega_n; \overbrace{\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}, \dots, \omega_{m_{\pi(l)}}}^l)}^{\overrightarrow{\rightarrow} \dots \overrightarrow{\rightarrow}} \\ &\quad \vdots \overbrace{\tilde{\mathbf{H}}(\omega_{m_{\pi(1)}}) \tilde{\mathbf{H}}(\omega_{m_{\pi(2)}}) \cdots \tilde{\mathbf{H}}(\omega_{m_{\pi(l)}})}^l + \dots \end{aligned} \quad (2.21)$$

可见其中所有 ω_{m_i} 均与式(2.1)中的 ω_{m_i} 完全相同，均是同一条光束 i 的角频率；相应地，所有以 ω_{m_i} 为自变量的因变量，也全都属于同一条光束 i ，比如该式中的 $\tilde{\mathbf{H}}(\omega_{m_i})$ 与同 i 的式(2.1)中的 $\tilde{\mathbf{E}}(\omega_{m_i})$ ，也均是同一条光束 i 的磁场和电场两条属性；同理，该式中的 $\tilde{\mathbf{M}}(\omega_n)$ 在生成的同时，也必伴随着式(2.1)中的 $\tilde{\mathbf{P}}(\omega_n)$ 同时生成(在其光路附近)，反之亦然；可见上式中的各阶非线性效应中的参与混频的磁场(频率)序列

$\{(\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}, \dots, \omega_{m_{\pi(l)}})\}$ 、简并因子，也与式(2.1)中相应阶磁场(频率)序列、简并因子分别相同。

同理，对于 $\tilde{\mathbf{H}}, \tilde{\mathbf{M}}, \tilde{\mathbf{B}}$ ，也有 $\tilde{\mathbf{H}}_n^* = \tilde{\mathbf{H}}_{-n}$ 、 $\tilde{\mathbf{M}}_n^* = \tilde{\mathbf{M}}_{-n}$ 、 $\tilde{\mathbf{B}}_n^* = \tilde{\mathbf{B}}_{-n}$ ，以及三个叠加式：

$$\tilde{\mathbf{H}}(\mathbf{r}, t) = \sum_{i=-\infty}^{\infty} \tilde{\mathbf{H}}_i(\mathbf{r}, t) \quad (2.22)$$

$$\begin{aligned} \tilde{\mathbf{M}}(\mathbf{r}, t) &= \sum_{n=-\infty}^{\infty} \tilde{\mathbf{M}}_n(\mathbf{r}, t) = \sum_{n=-\infty}^{\infty} \sum_{l=1}^{\infty} \tilde{\mathbf{M}}_n^{(l)}(\mathbf{r}, t) \\ &= \sum_{n=-\infty}^{\infty} e^{-i \sum_{i=1}^l \omega_i t} \sum_{l=1}^{\infty} \mathbf{M}_n^{(l)}(\mathbf{r}) = \sum_{n=-\infty}^{\infty} \mathbf{M}_n(\mathbf{r}) e^{-i \omega_n t} \end{aligned} \quad (2.23)$$

$$\begin{aligned} \tilde{\mathbf{B}}(\mathbf{r}, t) &= \sum_{n=-\infty}^{\infty} \tilde{\mathbf{B}}_n(\mathbf{r}, t) = \sum_{n=-\infty}^{\infty} \sum_{l=1}^{\infty} \tilde{\mathbf{B}}_n^{(l)}(\mathbf{r}, t) \\ &= \sum_{n=-\infty}^{\infty} e^{-i \sum_{i=1}^l \omega_i t} \sum_{l=1}^{\infty} \mathbf{B}_n^{(l)}(\mathbf{r}) = \sum_{n=-\infty}^{\infty} \mathbf{B}_n(\mathbf{r}) e^{-i \omega_n t} \end{aligned} \quad (2.24)$$

其中由于式(2.1)与式(2.21)的每一阶非线性效应均共享同一组参与混频的电磁场序列($\pi(1), \pi(2), \dots, \pi(l)$)，所以 $\tilde{\mathbf{M}}_n^{(l)}$ 与 $\tilde{\mathbf{B}}_n^{(l)} = \mu_0 \tilde{\mathbf{M}}_n^{(l)} (l \geq 2)$ 在合成过程中使用的原材料，是与 $\tilde{\mathbf{P}}_n^{(l)}$ 与 $\tilde{\mathbf{D}}_n^{(l)} = \tilde{\mathbf{P}}_n^{(l)} (l \geq 2)$ 所使用的原材料相同的 l 束波，的磁场部分($\tilde{\mathbf{H}}_1, \tilde{\mathbf{H}}_2, \dots, \tilde{\mathbf{H}}_l$)和电场部分($\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2, \dots, \tilde{\mathbf{E}}_l$)^①；因此不仅式(2.1)与式(2.21)的相同阶的非线性过程同时发生，且共享同一能量守恒过程 $\sum_{i=1}^l \omega_i = \omega_n$ 和动量守恒过程 $\sum_{i=1}^l \mathbf{k}_i = \mathbf{k}_n$ ；不仅同时合成出

^① 每一束波的 $\tilde{\mathbf{E}}_i$ 与 $\tilde{\mathbf{H}}_i$ 的波矢和频率相同，由四个麦氏方程程导出的电场、磁场两个波动方程的两个解保证。

$\tilde{\mathbf{M}}_n^{(l)}$ 与 $\tilde{\mathbf{P}}_n^{(l)}$ ，且二者的频率、波矢均相同^①；然后无数阶这样的非线性过程的结果之和 $\tilde{\mathbf{M}}_n^{NL}$ 与 $\tilde{\mathbf{P}}_n^{NL}$ （频率、波矢也均相同），再作为各自的波动方程右侧的非线性波源项，产生各自对应的和/差频波 $\tilde{\mathbf{H}}_n$ 与 $\tilde{\mathbf{E}}_n$ ，二者的频率共享 ω_n 且值仍为 $\sum_{i=1}^l \omega_i$ ，波矢也共享同一个 \mathbf{k}_n ，但值不一定为 $\sum_{i=1}^l \mathbf{k}_i$ ；接着 $\tilde{\mathbf{H}}_n$ 与 $\tilde{\mathbf{E}}_n$ 再通过一阶(非)线性效应，合成出 $\tilde{\mathbf{M}}_n^{(1)}$ 与 $\tilde{\mathbf{P}}_n^{(1)}$ 以及 $\tilde{\mathbf{B}}_n^{(1)} = \mu_0[\tilde{\mathbf{H}}_n + \tilde{\mathbf{M}}_n^{(1)}]$ 与 $\tilde{\mathbf{D}}_n^{(1)} = \epsilon_0 \tilde{\mathbf{E}}_n + \tilde{\mathbf{P}}_n^{(1)}$ ；最后合成出麦氏方程组中与 $\tilde{\mathbf{H}}_n$ 与 $\tilde{\mathbf{E}}_n$ 等价的 $\tilde{\mathbf{M}}_n = \tilde{\mathbf{M}}_n^{(1)} + \tilde{\mathbf{M}}_n^{NL}$ ， $\tilde{\mathbf{P}}_n = \tilde{\mathbf{P}}_n^{(1)} + \tilde{\mathbf{P}}_n^{NL}$ 、 $\tilde{\mathbf{B}}_n = \tilde{\mathbf{B}}_n^{(1)} + \tilde{\mathbf{B}}_n^{NL}$ ， $\tilde{\mathbf{D}}_n = \tilde{\mathbf{D}}_n^{(1)} + \tilde{\mathbf{D}}_n^{NL}$ ，至此麦氏方程组中六个物理量的所有阶次的值以及得来顺序都已大白。

2.1.1.2 三维空间中的含磁非线性电介质中的定态单色电场波动方程的最普遍形式

据前一节理论，将式(2.15)、式(2.16)中的关系 $\tilde{\mathbf{D}}_n = \tilde{\mathbf{D}}_n^{(1)} + \tilde{\mathbf{D}}_n^{NL} = \vec{\epsilon}_r^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}_n + \tilde{\mathbf{P}}_n^{NL}$ 代入单色化后的式(2.6)，并注意到 $\vec{\epsilon}_r^{(1)}$ 不含时，则式(2.6)变为

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}_n) + \frac{1}{c^2} \vec{\epsilon}_r^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{E}}_n}{\partial t^2} = -\frac{1}{\epsilon_0 c^2} \frac{\partial^2 \tilde{\mathbf{P}}_n^{NL}}{\partial t^2} - \frac{1}{\epsilon_0 c^2} \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n) \quad (2.25)$$

将式(2.20)中的 $\tilde{\mathbf{M}}_n = \tilde{\mathbf{M}}_n^{(1)} + \tilde{\mathbf{M}}_n^{NL}$ 及式(2.21)中的 $\tilde{\mathbf{M}}_n^{(1)} = \vec{\chi}_m^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}_n$ 代入上式得：

$$\frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n) = \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n^{(1)}) + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n^{NL}) \quad (2.26)$$

$$\begin{aligned} \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n^{(1)}) &= \frac{\partial}{\partial t} \left(\nabla \times [\vec{\chi}_m^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}_n] \right) \xrightarrow{\vec{\chi}_m^{(1)}(\omega_n) \text{不含时空}} \vec{\chi}_m^{(1)}(\omega_n) \cdot \frac{\partial}{\partial t} [\nabla \times \tilde{\mathbf{H}}_n] \\ &= \vec{\chi}_m^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{D}}_n}{\partial t^2} = \vec{\chi}_m^{(1)}(\omega_n) \cdot \frac{\partial^2 (\tilde{\mathbf{D}}_n^{(1)} + \tilde{\mathbf{D}}_n^{NL})}{\partial t^2} = \vec{\chi}_m^{(1)}(\omega_n) \cdot \frac{\partial^2 [\vec{\epsilon}_r^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}_n + \tilde{\mathbf{P}}_n^{NL}]}{\partial t^2} \\ &= \vec{\chi}_m^{(1)}(\omega_n) \cdot [\vec{\epsilon}_r^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{E}}_n}{\partial t^2} + \frac{\partial^2 \tilde{\mathbf{P}}_n^{NL}}{\partial t^2}] \end{aligned} \quad (2.27)$$

整理得非线性介质中单色电场波动方程的最普遍形式：

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}_n) + \vec{\mu}^{(1)}(\omega_n) \cdot \vec{\epsilon}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{E}}_n}{\partial t^2} = -\vec{\mu}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{P}}_n^{NL}}{\partial t^2} - \mu_0 \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n^{NL}) \quad (2.28)$$

由于该方程中各三维矢量均可分离出相同的时间部分 $e^{-i\omega_n t}$ ，即 $\tilde{\mathbf{E}}_n = \mathbf{E}_n(\mathbf{r}) e^{-i\omega_n t}$ 、

$\tilde{\mathbf{P}}_n^{NL} = \mathbf{P}_n^{NL}(\mathbf{r}) e^{-i\omega_n t}$ 、 $\tilde{\mathbf{M}}_n^{NL} = \mathbf{M}_n^{NL}(\mathbf{r}) e^{-i\omega_n t}$ ，因此可将该方程变为定态的：

$$\nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})] - \omega_n^2 \vec{\mu}^{(1)}(\omega_n) \cdot \vec{\epsilon}^{(1)}(\omega_n) \cdot \mathbf{E}_n(\mathbf{r})$$

^① 既由电磁的非线性过程原料角标和各角标的 $\tilde{\mathbf{E}}_i$ 与 $\tilde{\mathbf{H}}_i$ 的波矢和频率相同，也由麦氏方程组或波动方程约束。

$$= \omega_n^2 \overset{\rightarrow}{\mu}^{(1)}(\omega_n) \cdot \vec{P}_n^{NL}(\mathbf{r}) + \omega_n^2 \mu_0 [\nabla \times \vec{M}_n^{NL}(\mathbf{r})] \quad (2.29)$$

若忽略介质的非线性磁化 $\tilde{\mathbf{M}}_n^{NL}$, 只考虑线性磁化项 $\tilde{\mathbf{M}}_n^{(1)}$ 时, 上述方程变为:

$$\nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})] - \omega_n^2 \overset{\rightarrow}{\mu}^{(1)}(\omega_n) \cdot \overset{\rightarrow}{\epsilon}^{(1)}(\omega_n) \cdot \mathbf{E}_n(\mathbf{r}) = \omega_n^2 \overset{\rightarrow}{\mu}^{(1)}(\omega_n) \cdot \vec{P}_n^{NL}(\mathbf{r}) \quad (2.30)$$

若忽略介质的所有磁化 $\tilde{\mathbf{M}}_n$, 即介质为非磁介质, 为纯电介质时, 上述方程变为:

$$\nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})] - \frac{\omega_n^2}{c^2} \overset{\rightarrow}{\epsilon}^{(1)}(\omega_n) \cdot \mathbf{E}_n(\mathbf{r}) = \frac{\omega_n^2}{\epsilon_0 c^2} \vec{P}_n^{NL}(\mathbf{r}) \quad (2.31)$$

此即常见的三维空间中的无磁非线性电介质中的定态波动方程。

以上 2.1、2.1.1 章节可分别通过附录 C1.1、附录 C1.2 中 C1.2.1 至 C1.2.4 部分, 查看相应公式的具体推导过程。

2.1.2 磁各向同性的非线性电介质中的波动方程及其解

2.1.2.1 磁各向同性介质中, 完美相位匹配条件下, 波动方程的解随非线性效应的减弱而收敛到单色定振幅平面电磁波

在附录 C1.2.5 至附录 C1.2.10 部分, 详细验证了在 $\vec{P}_n^{NL}(\mathbf{r}) \approx \mathbf{0}$ 、完美相位匹配条件下, 磁各向同性介质中, 波动方程的解收敛于单色定振幅平面电磁波, 不论介质的一阶非线性系数(二阶极化率或介电常数张量)如何, 即无论介质是各向同性还是各向异性的、无论电矢量无源 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) = 0$ 是否成立; 也不关心介质的二阶以上非线性系数是否为零, 因为非线性效应或非线性极化毕竟相对于线性效应、线性极化来说, 确实太弱了, 即使转换效率比较高, 一阶和二阶非线性系数也不在同一个量级(见附录 C1.2.5.2 脚注部分), 因此波动方程的解相对于定振幅平面波变化不大。但这只是在方程的简化和解的近似上做这样的处理, 以使得对波动方程的解得以顺利进行下去, 然而实际计算非线性效应的时候, 因非线性源与方程解的性质直接相关而不能忽略。

2.1.2.2 各向异性非线性介质中波动方程的解的形式与各向同性非线性介质中的相同

由于已知磁各向同性非线性电介质中波动方程的解近似为定振幅平面电磁波, 而定振幅平面电磁波又是无源的, 所以 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) = 0$ 可直接代入式(C56)、式(C57)得到式(C58)。现将磁各向同性条件 $\mu_{nxz}^{(1)} = \mu_{nyy}^{(1)} = \mu_{nzz}^{(1)} = \mu_n^{(1)}$ 代入式(C58), 有:

$$\left[\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial x^2} + \omega_n^2 \mu_n^{(1)} \epsilon_{nxx}^{(1)} \right] E_{nx}(\mathbf{r}) = -\omega_n^2 \mu_n^{(1)} P_{nx}^{NL}(\mathbf{r})$$

$$\nabla^2 \mathbf{E}_n(\mathbf{r}) + \omega_n^2 \mu^{(1)}(\omega_n) \epsilon^{(1)}(\omega_n) \mathbf{E}_n(\mathbf{r}) = -\omega_n^2 \mu^{(1)}(\omega_n) \vec{P}_n^{NL}(\mathbf{r}) \quad (2.32)$$

但其中的电磁各向异性介质中的波动方程即使退化到磁各向同性, 因电各向异性

的存在，同样由附录 C1.2.5.4 给出的原因，也无法得出有物理意义的解，反而需要退回到电矢量有源的真实条件下，才可得出平面波解。因此这里只能利用后者，即电磁均各向同性介质中的波动方程，得出后者的解后，前者的解的形式因与后者的解的形式相同，而也算是迂回地得到了；而之所以电各向同性与电各向异性的非线性介质中波动方程的解的形式相同，需要如下证明一下：

由于在线性介质下，已经得到各向同性介质中的解与各向异性介质中的两个光束的解，共三个解的数学形式相同，均为定振幅平面电磁波；那么在非线性介质中，各向同性介质中的解，与各向异性介质中的两个光束的解，数学形式必将仍然相同，因为介质的各向同性、各向异性与介质的非线性无关——即介质的一阶非线性系数与高阶非线性系数不仅各分量取值无关，且 $\tilde{\mathbf{D}}_n, \tilde{\mathbf{P}}_n$ 的线性部分 $\tilde{\mathbf{D}}_n^{(1)} = \epsilon_0 \tilde{\mathbf{E}}_n + \tilde{\mathbf{P}}_n^{(1)}$, $\tilde{\mathbf{P}}_n^{(1)}$ 均在电各向同性和各向异性的两个波动方程左侧，作为 $\tilde{\mathbf{E}}_n$ 的一次项的构成部分之一，而 $\tilde{\mathbf{D}}_n, \tilde{\mathbf{P}}_n$ 的非线性部分 $\tilde{\mathbf{D}}_n^{NL} = \tilde{\mathbf{P}}_n^{NL}$ 则均位于波动方程右侧，作为非线性波源项即“常数项”（同时也是 $\tilde{\mathbf{E}}_n$ 的高次项）存在，即二者对波动方程的影响也是无交集的，所以二者对波动方程的解的影响也是无交集的、互不干扰的、相互平行的。

从附录 C 知，线性部分对波动方程的解的影响是，从各向同性介质到各向异性介质会出来两个折射率曲面，导致可能的介电常数数目从一个变为两个，进而导致双波动方程和两个解，其影响是增加波动方程和解的数量；从接下来的过程可知，波动方程非线性部分对波动方程的解的影响是，从线性介质到非线性介质，解中的定振幅会变成随空间缓变的振幅。

因此当“线性介质中所有波动方程的解的数学形式统一”这一结论成立时，非线性介质中同样有：各向同性的非线性介质的波动方程的解的形式，必然也是各向异性非线性介质的波动方程的解的形式，且是各向异性的非线性介质中 o 光和 e 光公有的解的形式。因此接下来的内容，适用于铌酸锂晶体这一既是非线性介质、同时也是各向异性介质的负单轴晶体。即下面的部分可直接用各向同性的非线性介质中的波动方程即式(2.32)的二式，毕竟其解的形式与式(2.32)的一式的两个解的形式是相同的。

2.1.2.3 描述和频波生成过程的耦合波方程

（1）特定几何状态下铌酸锂晶体中的二阶和频极化光

当式(2.32)右侧非线性项为零时，波动方程退化为电磁各向同性介质中的波动方程，其解为平面电磁波复振幅 $\mathbf{E}_n(\mathbf{r}) = \mathbf{A}_n e^{ik_n \cdot \mathbf{r}}$ ，其中当一阶非线性系数是实数时，

$\mathbf{A}_n = \mathbf{E}_{n0}$ 为实的定振幅，当一阶非线性系数为复数时， $\mathbf{A}_n(\mathbf{r}) = \mathbf{E}_{n0} \cdot e^{-\alpha_n \cdot \mathbf{r}}$ 因含有衰减因

予 $e^{-\alpha_n \cdot r}$ 而含空^①。而受到解线性微分方程的常数变易法的启发，对于此二阶非线性微分方程，当一、二阶非线性系数为实数时，也假设其通解在形式上为 $\mathbf{E}_n(\mathbf{r}) = \mathbf{A}_n(\mathbf{r})e^{ik_n \cdot r}$ ，其相位部分与平面电磁波相同，但振幅部分为含空的实的振幅。当非线性系数不大时，振幅是缓变的；当非线性系数趋于零时，缓变振幅趋近定振幅。

现只考察二阶非线性效应中的 SFG(Sum-Frequency Generation)过程对应的和频极化光作为波源所生成的和频波解，则式(2.32)右侧非线性波源项 $\mathbf{P}_n^{NL}(\mathbf{r})$ 只剩 $\mathbf{P}_n^{(2)}(\mathbf{r})$ ：

$$\nabla^2 \mathbf{E}_n(\mathbf{r}) + \omega_n^2 \mu^{(1)}(\omega_n) \varepsilon^{(1)}(\omega_n) \mathbf{E}_n(\mathbf{r}) = -\omega_n^2 \mu^{(1)}(\omega_n) \mathbf{P}_n^{(2)}(\mathbf{r}) \quad (2.33)$$

现考察其右侧非线性波源项 $\mathbf{P}_n^{(2)}(\mathbf{r})$ ，据附录 E 的式(E37)，在将二阶非线性系数通过克莱曼对称性化简到独立分量个数只剩 10 个后，对于两束给定的频率不同的参与混频的单色光场 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ ，二者通过二阶非线性过程极化得的和频极化光复振幅为：

$$\begin{bmatrix} P_x^{(2)}(\omega_1 + \omega_2) \\ P_y^{(2)}(\omega_1 + \omega_2) \\ P_z^{(2)}(\omega_1 + \omega_2) \end{bmatrix} = 4\epsilon_0 \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{16} & d_{22} & d_{23} & d_{24} & d_{14} & d_{12} \\ d_{15} & d_{24} & d_{33} & d_{23} & d_{13} & d_{14} \end{bmatrix} \cdot \begin{bmatrix} E_x(\omega_1)E_x(\omega_2) \\ E_y(\omega_1)E_y(\omega_2) \\ E_z(\omega_1)E_z(\omega_2) \\ E_y(\omega_1)E_z(\omega_2) + E_z(\omega_1)E_y(\omega_2) \\ E_x(\omega_1)E_z(\omega_2) + E_z(\omega_1)E_x(\omega_2) \\ E_x(\omega_1)E_y(\omega_2) + E_y(\omega_1)E_x(\omega_2) \end{bmatrix} \quad (2.34)$$

由于铌酸锂 LiNbO₃ 晶体属于三方晶系(Trigonal System)，查表可知三方晶系的二阶非线性系数在当光轴平行于三方晶系不同的 n 度旋转轴、旋转反演轴等点群所对应的对称轴时，非零分量数和种类也不同：

表 2.1 笛卡尔坐标系下三方晶系的二阶非线性系数张量的非零分量及其关系式

Crystal System	Crystal Class	Nonvanishing Tensor Elements
Trigonal	$3 = C_3$	$xxx = -xyy = -yyz = -yxz, xyz = -yxz, xzy = -yzx,$ $xzx = yzy, xxz = yyz, yyy = -yxz = -xxy = -xyx,$ $zxx = zyy, zzz, zxy = -zyx$
	$32 = D_3$	$xxx = -xyy = -yyx = -yxy, xyz = -yxz,$ $xzy = -yzx, zxy = -zyx$
	$3m = C_{3v}$	$xzx = yzy, xxz = yyz, zxx = zyy, zzz, yyy = -yxz = -xxy = -xyx$ (mirror plane perpendicular to \hat{x})
	$\bar{3} = S_6, \bar{3}m = D_{3d}$	Each element vanishes

由于即使在克莱曼对称性下，以及再加上额外的属于中级晶族的三方晶系的固有宏观对称性，二阶非线性系数张量分量仍较多，且不为零的分量也较多，为尽可能地

^① 此时复振中相位部分也会从 $e^{ik_n \cdot r}$ 变为 $e^{i\beta_n \cdot r}$ ， \mathbf{k}_n 从实波矢变为复波矢 $\mathbf{k}_n = \boldsymbol{\beta}_n + i\boldsymbol{\alpha}_n$ 。

少引入方程右侧非线性效应的系数及物理量的各分量，需要考虑和构建一系列特殊场景的组合：晶体外入射光场的入射方向、晶体内混频波的传播方向、极化方向，以及晶体、晶面相对于入射光的姿势放置情况。

现考虑单色平行连续入射光垂直晶面地入射无损的铌酸锂非线性光学介质，并忽略铌酸锂晶体中 o、e 光的双折射现象（波动方程的标量介电常数在之后可替换为铌酸锂单轴晶体的 o 光或 e 光的等效标量介电常数）；将该几何场景和所选坐标系下的三维铌酸锂晶体的二阶非线性系数的各分量代入式(2.34)中，并两边取大小（注意不是取模，要保留场的相位部分），得到 SFG 对应的标量极化光的复振幅表达式：

$$P^{(2)}(\omega_3) = 4\epsilon_0 d_{eff} \cdot E(\omega_1)E(\omega_2) \quad (2.35)$$

其中 $P^{(2)}(\omega_3) = |\mathbf{P}^{(2)}(\omega_3)| \arg[\mathbf{P}^{(2)}(\omega_3)]$ 、 $E(\omega_1) = |\mathbf{E}(\omega_1)| \arg[\mathbf{E}(\omega_1)]$ 、 $E(\omega_2) = |\mathbf{E}(\omega_2)| \arg[\mathbf{E}(\omega_2)]$ 。各大晶系的 d_{eff} 的计算工作已由 Midwinter and Warner (1965)，以及 Zernike and Midwinter (1973) 给出，对于其中三方晶系中的负单轴晶体，且点群为 $3m$ 的复三方单锥晶类晶体，有：

$$d_{eff} = d_{31} \sin \theta - d_{22} \cos \theta \sin 3\phi \quad (2.36)$$

其中假设克莱曼对称性的前提成立，即介质的色散可以忽略，以至 d_{eff} 不含 ω_3 。除了幅值和相位部分外，被产生的 $\mathbf{P}^{(2)}(\omega_3)$ 的极化方向相对于产生它的 $\mathbf{E}(\omega_1)$ 、 $\mathbf{E}(\omega_2)$ 偏振方向的相对方位，也需通过式(2.34)确定，因而一般并不共线更不太可能同向，这一点与“一阶极化光 $\mathbf{P}^{(1)}(\omega_3)$ 的极化方向与产生其线性极化的 $\mathbf{E}(\omega_3)$ 的偏振方向不同”(C128)类似，但二阶情况下的偏离程度(并非离散角)及其不可预测性即计算的复杂程度可能都要更胜一筹。甚至连 $\mathbf{P}^{(2)}(\omega_3)$ 相对于 $\mathbf{E}(\omega_1), \mathbf{E}(\omega_2)$ 的大小也不好确定，因其虽受非线性系数较小的限制，但正比于入射光场的二次方。

(2) 和频极化光 $\mathbf{P}^{(2)}(\omega_3)$ 的极化方向与和频光 $\mathbf{E}(\omega_3)$ 的偏振方向大致相同，非线性全息效率不会显著下降

而若考虑所产生的和频光 $\mathbf{E}(\omega_3)$ 相对于产生它的和频极化波源 $\mathbf{P}^{(2)}(\omega_3)$ 的极化方向，则不应考察线性或非线性的极化过程，而应考察波动方程，并且不能考察接下来标量情形下的波动方程，也不能考察式(2.33)这样的虽为矢量形式，但做了无源近似的波动方程，而应考察代入了二阶非线性效应的不做无源近似的矢量形式波动方程如(C62)，或考察波矢形式的麦氏方程组所导出的(C109)与(C113)：由于之后会解出试探解的振幅 $A_3(z)$ 的大小 $A_3(z)$ 确实随传播方向上的位移而缓变，因此不仅 $\mathbf{E}(\omega_3)$ 所参与

的二阶及以上的和频极化过程所生成的 $\mathbf{P}^{(l \geq 2)}(\omega_3 + \dots) = \mathbf{D}^{(l \geq 2)}(\omega_3 + \dots)$ 的振幅是缓变的，就连 $\mathbf{E}(\omega_3)$ 在生成后，于介质中单独线性极化出的 $\mathbf{P}^{(1)}(\omega_3), \mathbf{D}^{(1)}(\omega_3)$ ，在波矢匹配结束之前，也均是缓变振幅的。

因此考虑在晶体中单独存在的 $\mathbf{E}(\omega_3)$ ，它在产生之后即使与其他频率的混频光一起继续合成高阶非线性极化光^[29]，或者在满足波矢匹配条件时自己与自己一起作为原料产生倍频极化光等，铌酸锂晶体中频率为 ω_3 的光仍只有产生 $\mathbf{E}(\omega_3)$ 的二阶非线性波源 $\mathbf{P}^{(2)}(\omega_3)$ 以及 $\mathbf{E}(\omega_3)$ ，其中假定了产生二阶和频极化光 $\mathbf{P}^{(2)}(\omega_3)$ 的混频光的来源只有 $\mathbf{E}(\omega_1), \mathbf{E}(\omega_2)$ ，没有其他频率的两光参与类似的 SFG 过程，并且假定没有更多频率的光参与三阶及以上的 SFG 过程以合成出 $\mathbf{E}(\omega_3)$ 的子波，此时频率为 ω_3 且可出射晶体的非极化光，只包含唯一一个以 $\mathbf{P}^{(2)}(\omega_3)$ 为非线性波源的波动方程的解 $\mathbf{E}(\omega_3)$ ，而不是多个波动方程的解的集合。那么在这样的情况下，单一 $\mathbf{E}(\omega_3)$ 在生成后，即使它在非线性介质中，后续也只可能发生线性极化，并极化出 $\mathbf{P}^{(1)}(\omega_3), \mathbf{D}^{(1)}(\omega_3)$ 。——因此此时晶体中频率为 ω_3 的光至多同时至少，有三种，一是由 $\mathbf{E}(\omega_1), \mathbf{E}(\omega_2)$ 非线性极化合成的 $\mathbf{P}^{(2)}(\omega_3) = \mathbf{D}^{(2)}(\omega_3)$ ，二是 $\mathbf{P}^{(2)}(\omega_3)$ 通过波动方程生成的 $\mathbf{E}(\omega_3)$ ，三是 $\mathbf{E}(\omega_3)$ 所线性极化出的 $\mathbf{P}^{(1)}(\omega_3), \mathbf{D}^{(1)}(\omega_3)$ 。其中只有第一者 $\mathbf{P}^{(2)}(\omega_3) = \mathbf{D}^{(2)}(\omega_3)$ 是相对更定振幅的（非线性效率不高时），后两者 $\mathbf{E}(\omega_3), \mathbf{P}^{(1)}(\omega_3), \mathbf{D}^{(1)}(\omega_3)$ 均是缓变振幅的，但当波矢匹配过程结束后，后两者就变为定振幅的了。

那么式(C94)中的 $\mathbf{E}_n = \mathbf{E}_3, \mathbf{D}_n = \mathbf{D}_3 = \mathbf{D}_3^{(1)} + \mathbf{D}_3^{(2)}$ ^①，在波矢匹配过程结束前，由于

^① 即使 $\mathbf{D}^{(2)}(\omega_3)$ 的来源不是来自 $\mathbf{E}(\omega_3)$ 而来自 $\mathbf{E}(\omega_1), \mathbf{E}(\omega_2)$ ，由于单色麦氏方程组只关心生成的所有同频单色波，考虑的是结果，而不考察这些结果的得来源泉和得来过程，所以 $\mathbf{D}^{(2)}(\omega_3)$ 并不在 $\mathbf{E}(\omega_1), \mathbf{E}(\omega_2)$ 所在的单色麦氏方程组中，而在与其同频的 $\mathbf{E}(\omega_3)$ 所在的单色麦氏方程组中；即使不是这样，那么也不清楚 $\mathbf{D}^{(2)}(\omega_3)$ 到底属于 $\mathbf{E}(\omega_1)$ 的极化还是属于 $\mathbf{E}(\omega_2)$ 的极化，因为 $\mathbf{D}^{(2)}(\omega_3)$ 是二者共同的二阶非线性极化，因此既不能放在 $\mathbf{E}(\omega_1)$ 所在的单色麦氏方程组中，又不能放在 $\mathbf{E}(\omega_2)$ 所在的单色麦氏方程组中，所以 $\mathbf{D}^{(2)}(\omega_3)$ 除了能与 $\mathbf{E}(\omega_3)$ 一对一匹配之外，别无它处可收留。而若即使 $\mathbf{D}^{(2)}(\omega_3)$ 确实不在 $\mathbf{E}(\omega_3)$ 所在的单色麦氏方程组中，以至 $\mathbf{D}_3 = \mathbf{D}_3^{(1)}$ ，则此时 \mathbf{D}_3 也是含空的。 $\mathbf{P}^{(2)}(\omega_3) = \mathbf{D}^{(2)}(\omega_3)$ 如此奇特，地位如此尴尬，怪不得可作为非线性全息的参考光和重现光——虽然从其频率与 $\mathbf{E}(\omega_3)$ 恒等来看它确实属于 $\mathbf{E}(\omega_3)$ ，但从其波矢与 $\mathbf{E}(\omega_3)$ 可能不等而恒等于 $\mathbf{E}(\omega_1), \mathbf{E}(\omega_2)$ 的波矢之和看来，它又属于 $\mathbf{E}(\omega_1), \mathbf{E}(\omega_2)$ ，并且由于波矢不同而传播方向不同而与 $\mathbf{E}(\omega_3)$ 分属两束光，以至于可以作为非线性全息的参考光和重现光（当然具体的光线而非相位传播方向要看能流密度矢量）。那么归根结底它到底属于谁，还是均独立于二者，就要看这里的实验了。该思维实验转换为实际实验后，其非零结果可证明 $\mathbf{P}^{(2)}(\omega_3) = \mathbf{D}^{(2)}(\omega_3)$ 属于 $\mathbf{E}(\omega_3)$ ，零则不属于。

其中 \mathbf{E}_3 与 $\mathbf{D}_3^{(1)}$ 的振幅均随空间缓变，因此上述物理量的旋度和散度在分别转换为波矢与它们的叉积时，等号会变为约等号，以至于 \mathbf{D}_3 只是近似 $\pm \mathbf{k}_3$ ，且式(C94)、(C109)、(C113)的等号均变为约等号，所得的离散角(C128)相对于定振幅的 $\mathbf{E}_3, \mathbf{D}_3$ 间的离散角有些许变化，但变化不大。因此若要考慮 $\mathbf{P}_3^{(2)}$ 相对于 \mathbf{E}_3 的方向的一种可行的办法是，先假定 $\mathbf{E}_3, \mathbf{D}_3$ 均为定振幅，通过代入含有 \mathbf{E}_3 到 \mathbf{D}_3 的线性和非线性极化关系的新的波矢形式的麦氏方程组的三号方程，所导出的新的离散角公式，算出 $\mathbf{D}_3^{(1)} + \mathbf{D}_3^{(2)} = \mathbf{D}_3$ 相对于 \mathbf{E}_3 的大致方向，再通过线性极化式 $\mathbf{D}_3^{(1)} = \overleftrightarrow{\epsilon}_3^{(1)} \cdot \mathbf{E}_3$ 算出相对于 $\mathbf{D}_3^{(1)}$ 相对于 \mathbf{E}_3 的准确方向，然后二者作差即可得到 $\mathbf{P}^{(2)}(\omega_3) = \mathbf{D}^{(2)}(\omega_3)$ 相对于 \mathbf{E}_3 的大致方向，其中的误差只在于微分形式的麦氏方程组转换为波矢形式的麦氏方程组时，假定 $\mathbf{E}_3, \mathbf{D}_3$ 为定振幅而实际却是缓变振幅所带来的误差。而如果考察波矢匹配过程结束后的情况，则因振幅确定下来了而没有误差。

上述过程中之所以需要考慮 $\mathbf{P}_3^{(2)}$ 相对于 \mathbf{E}_3 的方向，是因正空间非线性全息中的参考光和重现光均是 $\mathbf{P}_3^{(2)}$ ，在晶体内与 \mathbf{E}_3 发生干涉生成干涉图样，以及照射干涉图样生成还原的倍频物光即产生重建像的重现光，均是 $\mathbf{P}_3^{(2)}$ ；而如果 $\mathbf{P}_3^{(2)}$ 的极化方向与 \mathbf{E}_3 的偏振方向的夹角过大，则由于二者相互垂直的偏振分量不存在相消和相长，叠加后只是改变了偏振方向，对干涉条纹的强弱分布没有贡献，将导致参与干涉的有效振幅降低，以至生成的全息图整体强度下降，并且导致再现光生成重建像过程中的衍射效率也不高。所幸的是，特定几何的实验上 $\mathbf{P}_3^{(2)}, \mathbf{E}_3$ 的振动方向偏离并不严重，毕竟作为非线性源的 $\mathbf{P}^{(2)}(\omega_3)$ ，其所产生的和频光 $\mathbf{E}(\omega_3)$ 的偏振方向应该与 $\mathbf{P}^{(2)}(\omega_3)$ 其本身的极化方向大致同向（在之后会通过二者均与 $\mathbf{E}(\omega_1), \mathbf{E}(\omega_2)$ 大致同向以迂回地得到证明），且二者的波矢方向在傍轴近似下近似同向，而波矢大小只在完美相位匹配条件满足时，才相同；二者的大小也不好比较，因之后会给出一维条件下 $A_3(z) \propto d_{\text{eff}} \cdot E_{01} E_{02} \cdot z$ ，不仅正比于入射场的二次方与非线性系数之积，还在完美相位匹配和入射光场定振幅情况下正比于传播距离，因此参与混频的入射光、极化光、和频光三者的相对大小连同相对方向均不好一概而论，需视情况而定。总之，特定几何下，由于 $\mathbf{P}_3^{(2)}$ 与 \mathbf{E}_3 频率相同、振动方向相近、尽管波矢不同但因频率相同而在相同空间点处相位差恒定，再加上振幅大小相近、光程差相差不大的附加条件近似得到满足，二者便可发生稳定的干涉形成干涉图样，正空间非线性全息效率并不会降低太多。

(3) 描述同一个三波混频过程的三个耦合波方程

现假设 $\tilde{\mathbf{E}}_1 = A_1 e^{i(k_1 \cdot r - \omega_1 t)} = \mathbf{E}_{01} e^{i(k_1 z - \omega_1 t)}$, $\tilde{\mathbf{E}}_2 = A_2 e^{i(k_2 \cdot r - \omega_2 t)} = \mathbf{E}_{02} e^{i(k_2 z - \omega_2 t)}$ 以及 $\tilde{\mathbf{P}}_3^{(2)} =$

$\mathbf{P}_{03}^{(2)} e^{i[(k_1+k_2) \cdot r - (\omega_1 + \omega_2)t]} = \mathbf{P}_{03}^{(2)} e^{i[(k_1+k_2)z - \omega_3 t]}$ 三者的传播方向均沿 z 轴正向（注意 z 轴在这里不一定为光轴方向；尽管在各阶非线性系数中的下角标中的 z 均指代光轴方向的分量，但那里的 z 与这里的 z 不同），则式(2.35)写作：

$$P_3^{(2)}(z) = 4\epsilon_0 d_{eff} \cdot E_1(z)E_2(z) = 4\epsilon_0 d_{eff} \cdot E_{01}E_{02}e^{i(k_1+k_2)z} = P_{03}^{(2)} e^{i(k_1+k_2)z} \quad (2.37)$$

同时波动方程(2.33)写作标量形式：

$$\frac{\partial^2 E_3(z)}{\partial z^2} + \omega_3^2 \mu_3^{(1)} \epsilon_3^{(1)} E_3(z) = -\omega_3^2 \mu_3^{(1)} P_3^{(2)}(z) \quad (2.38)$$

其中由于在将该波动方程化为定态之初，就已经假设波动方程的解

$\tilde{\mathbf{E}}_3 = \mathbf{A}_3(\mathbf{r})e^{i(\mathbf{k}_3 \cdot \mathbf{r} - \omega_3 t)}$ 与 $\tilde{\mathbf{P}}_3^{(2)} = \mathbf{P}_{03}^{(2)} e^{i[(k_1+k_2)z - \omega_3 t]}$ 共享同一个频率，即 $\omega_3 = \omega_1 + \omega_2$ ；但并不一定共享同一个波矢，即并不一定有 $\mathbf{k}_3 = \mathbf{k}_1 + \mathbf{k}_2$ 、 $k_3 = k_1 + k_2$ ；但认为有 $\tilde{\mathbf{E}}_3$ 的波矢 \mathbf{k}_3 与 $\tilde{\mathbf{P}}_3^{(2)}$ 的波矢 $\mathbf{k}_1 + \mathbf{k}_2$ 大致同向，之后可知三维空间中二者方向的偏离程度由波矢失配矢量和相应倒格矢决定；然而振动方向不一定同向。可见 $\tilde{\mathbf{P}}_3^{(2)}$ 的角标 3 仅仅是在暗示其 $\mathbf{P}^{(2)}(\omega_3)$ 频率为 ω_3 ，但其波矢并不为 \mathbf{k}_3 ，因此不能将 $\tilde{\mathbf{P}}_3^{(2)}$ 写作 $\mathbf{P}^{(2)}(\omega_3, \mathbf{k}_3)$ 或 $\mathbf{P}^{(2)}(\mathbf{k}_3)$ ，且 $\mathbf{E}(\omega_3), \mathbf{P}^{(1)}(\omega_3), \mathbf{D}^{(1)}(\omega_3)$ 不仅振幅含空缓变，其波矢 \mathbf{k}_3 还不一定为 $\mathbf{k}_1 + \mathbf{k}_2$ ，而 $\mathbf{P}^{(2)}(\omega_3) = \mathbf{D}^{(2)}(\omega_3)$ 不仅振幅是定矢量，其波矢还一定为 $\mathbf{k}_1 + \mathbf{k}_2$ ，但尽管二个阵营在振幅和波矢上有差异，它们却有共同的频率 $\omega_3 = \omega_1 + \omega_2$ 。

因 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 均沿 z 轴正向传播， $\mathbf{E}_3(\mathbf{r}) = \mathbf{A}_3(\mathbf{r})e^{ik_3 \cdot \mathbf{r}}$ 的波矢将与 $\mathbf{k}_1 + \mathbf{k}_2$ 、 \mathbf{k} 完全同向，于是可将其标量形式写作：

$$E_3(z) = A_3(z)e^{ik_3 z} \quad (2.39)$$

将式(2.39)与(2.37)代入波动方程(2.38)，并注意到据 $v_3^2 = 1/\left[\mu_3^{(1)} \epsilon_3^{(1)}\right] = \omega_3^2/k_3^2$ 有 $k_3^2 = \omega_3^2 \mu_3^{(1)} \epsilon_3^{(1)}$ ，以至下面一式中左侧第三项与第四项抵消，得到：

$$\left[\frac{d^2 A_3(z)}{dz^2} e^{ik_3 z} + 2 \frac{dA_3(z)}{dz} ik_3 e^{ik_3 z} - k_3^2 A_3(z) e^{ik_3 z} \right] + \omega_3^2 \mu_3^{(1)} \epsilon_3^{(1)} A_3(z) e^{ik_3 z} = -\omega_3^2 \mu_3^{(1)} P_{03}^{(2)} e^{i(k_1+k_2)z}$$

$$\frac{d^2 A_3(z)}{dz^2} + 2ik_3 \frac{dA_3(z)}{dz} = -\frac{k_3^2}{\epsilon_3^{(1)}} P_{03}^{(2)} e^{i(k_1+k_2-k_3)z} \quad (2.40)$$

当 $P_3^{(2)}(z)$ 很小时，可以采用缓变振幅近似，即认为

$$\left| \frac{d^2 A_3(z)}{dz^2} \right| \ll \left| k_3 \frac{dA_3(z)}{dz} \right| \quad (2.41)$$

于是方程(2.40)第一项可忽略，得到耦合振幅方程：

$$\frac{dA_3(z)}{dz} = \frac{ik_3}{2\epsilon_3^{(1)}} P_{03}^{(2)} e^{i\Delta k \cdot z} \xrightarrow{\frac{k_3 |\omega_3| = |\omega_3 n_3| = |\omega_3| n_3}{\epsilon_3^{(1)} = \epsilon_0 \epsilon_{r3}^{(1)} = \epsilon_0 n_{3e} |n_{3e}| = \epsilon_0 n_{3e}^2}, \frac{|\omega_3| \sqrt{\mu_{r3}^{(1)} c_{r3}^{(1)}}}{c} = \frac{|\omega_3| \sqrt{\mu_{r3}^{(1)} n_{3e}}}{c}} \frac{i |\omega_3| \sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c} P_{03}^{(2)} e^{i\Delta k \cdot z} \quad (2.42)$$

其中定义了波矢失配量 $\Delta k = k_1 + k_2 - k_3$ 、一阶非线性系数和等效标量折射率默认为正、但即使是和频，频率 $\omega_3 = \omega_1 + \omega_2$ 也不一定为正；一般的无磁非线性电介质中 $\mu_m^{(1)} = 0$ ，此时 $\sqrt{\mu_{r3}^{(1)}} = 0$ 。该方程描绘了三波混频过程中，目标合成光的振幅 $A_3(z)$ 与参与合成的两束光的场振幅 $A_1 = E_{01}$ 、 $A_2 = E_{02}$ 间的关系。

现若参与混频的两束原料光 $\tilde{E}_1 = A_1(z) e^{i(k_1 z - \omega_1 t)}$ ， $\tilde{E}_2 = A_2(z) e^{i(k_2 z - \omega_2 t)}$ 的振幅也含空，则由二者极化出的和频极化光 $\tilde{P}_3^{(2)} = p_3^{(2)}(z) e^{i[(k_1 + k_2)z - \omega_3 t]}$ 的振幅也含空，但此时耦合振幅方程(2.42)仍成立，因为对右侧非线性波源项没有求导，整个推导过程保持不变：

$$\frac{dA_3(z)}{dz} = \frac{ik_3}{2\epsilon_3^{(1)}} p_3^{(2)}(z) e^{i\Delta k \cdot z} = \frac{i |\omega_3| \sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c} p_3^{(2)}(z) e^{i\Delta k \cdot z} \quad (2.43)$$

该扩展后的耦合振幅方程将 ω_1, ω_2 两束光的振幅变化也考虑了进去，并将 ω_1, ω_2 从原料光的地位升格为了与 $\omega_3 = \omega_1 + \omega_2$ 相等的地位，一方面诠释了“2→1”这一两束光到第三束光的合成过程，实际上是三束光的混频过程，即“三波混频”过程，三束光的地位是等价的；另一方面解锁了在满足相位匹配的条件下， ω_3 还可与 ω_1, ω_2 等光继续参与二阶或高阶混频，继续混频出更高频、波矢更长、但振幅更小且随空间变化更剧烈的和频光，此时 ω_3 从被合成的身份，转变为参与合成的原料光即原 ω_1, ω_2 的角色，而由于该 ω_3 光已经振幅含空、波矢不一定为 $k_1 + k_2$ 了，以至于它所合成的更高阶更高频的和频光，其振幅将“更加含空”、波矢也更不一定为 $k_1 + k_2 + \dots$ 了；同时，这也说明二到一的“合成过程”，实际上应被看做三束光中其中任意 3-1 束光，到另 1 束光的“转换过程”，因此一共有 $A_3^1 = 3$ 个转换过程，而 $\omega_3 = \omega_1 + \omega_2$ 只是其中之一，还应有 $\omega_1 = \omega_3 - \omega_2$ 、 $\omega_2 = \omega_3 - \omega_1$ 所分别对应的转换过程：

$$\frac{\partial^2 E_1(z)}{\partial z^2} + \omega_1^2 \mu_1^{(1)} \epsilon_1^{(1)} E_1(z) = -\omega_1^2 \mu_1^{(1)} P_1^{(2)}(z) = -\omega_1^2 \mu_1^{(1)} p_1^{(2)}(z) e^{i(k_3 - k_2)z} \quad (2.44)$$

$$\frac{\partial^2 E_2(z)}{\partial z^2} + \omega_2^2 \mu_2^{(1)} \epsilon_2^{(1)} E_2(z) = -\omega_2^2 \mu_2^{(1)} P_2^{(2)}(z) = -\omega_2^2 \mu_2^{(1)} p_2^{(2)}(z) e^{i(k_3 - k_1)z} \quad (2.45)$$

其中据类似式(2.34-2.37)的过程，有

$$\begin{aligned} P_1^{(2)}(z) &= P^{(2)}(\omega_1, k_1) = 4\epsilon_0 d_{eff}(\omega_1) \cdot E(\omega_3, k_3) E(-\omega_2, -k_2) = 4\epsilon_0 d_{eff}(\omega_1) \cdot E(\omega_3, k_3) E^*(\omega_2, k_2) \\ &= 4\epsilon_0 d_{eff}(\omega_1) \cdot E_3(z) E_2^*(z) = 4\epsilon_0 d_{eff}(\omega_1) \cdot A_3(z) A_2^*(z) e^{i(k_3 - k_2)z} =: p_1^{(2)}(z) e^{i(k_3 - k_2)z} \end{aligned} \quad (2.46)$$

$$P_2^{(2)}(z) = 4\epsilon_0 d_{eff}(\omega_2) \cdot A_3(z) A_1^*(z) e^{i(k_3 - k_1)z} =: p_2^{(2)}(z) e^{i(k_3 - k_1)z} \quad (2.47)$$

其中由于克莱曼对称性，有 $d_{u_1 u_{23}}(\omega_1) = d_{u_1 u_{23}}(\omega_2) = d_{u_1 u_{23}}(\omega_3)$ ，将其代入式(2.36)得：

$$d_{eff}(\omega_1) = d_{eff}(\omega_2) = d_{eff}(\omega_3) = d_{eff} \quad (2.48)$$

因此代入得：

$$\begin{aligned} p_1^{(2)}(z) &= 4\epsilon_0 d_{eff} \cdot A_3(z) A_2^*(z) \\ p_2^{(2)}(z) &= 4\epsilon_0 d_{eff} \cdot A_3(z) A_1^*(z) \end{aligned} \quad (2.49)$$

再将式(2.49)代入式(2.44-2.45)，并通过类似式(2.40-2.43)的过程，得类似(2.43)的两个额外的耦合振幅方程：

$$\frac{dA_1(z)}{dz} = \frac{ik_1}{2\epsilon_1^{(1)}} p_1^{(2)}(z) e^{-i\Delta k \cdot z} = \frac{i|\omega_1| \sqrt{\mu_{r1}^{(1)}}}{2\epsilon_0 n_{1e} c} p_1^{(2)}(z) e^{-i\Delta k \cdot z} \quad (2.50)$$

$$\frac{dA_2(z)}{dz} = \frac{ik_2}{2\epsilon_2^{(1)}} p_2^{(2)}(z) e^{-i\Delta k \cdot z} = \frac{i|\omega_2| \sqrt{\mu_{r2}^{(1)}}}{2\epsilon_0 n_{2e} c} p_2^{(2)}(z) e^{-i\Delta k \cdot z} \quad (2.51)$$

式(2.43)、(2.50)、(2.51)三个耦合振幅方程一起共同描述同一个三波混频过程。

2.1.2.4 相位匹配问题

(1) 完美相位匹配时，和频波振幅随传播距离线性增长

当参与混频的两个入射光场到和频光的转换不充分、转换效率不高时，和频光 \tilde{E}_3 从 \tilde{E}_1, \tilde{E}_2 中吸取的能量并不多，不会造成 \tilde{E}_1, \tilde{E}_2 能量和振幅的显著下降，此时两个参与混频的入射光场振幅近似恒定，式(2.43)中的 $A_1(z) = E_{01}, A_2(z) = E_{02}$ 不含空，则式(2.43)退化到式(2.42)，此时若再有波矢失配为零，即

$$\Delta k = k_1 + k_2 - k_3 = 0 \quad (2.52)$$

则将式(2.42)两边对 z 积分得：

$$A_3(z) = \frac{i|\omega_3| \sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c} P_{03}^{(2)} \cdot z \quad (2.53)$$

此时和频光的振幅的模 $|A_3(z)|$ 随着和频光传播距离 z 而线性增大，这首先说明该解符合缓变振幅条件(2.41)，反过来说明缓变振幅近似条件是正确的、符合现实的、与其所导致的方程解是自洽的。下面考虑其物理意义：从波的角度，此时和频光 $\tilde{E}_3 = A_3(z) e^{i(k_3 z - \omega_3 t)}$ 与二阶非线性和频极化光 $\tilde{P}_3^{(2)} = P_{03}^{(2)} e^{i[(k_1 + k_2)z - \omega_3 t]}$ 的波矢和频率都相同，因此它们的相位差恒定，于是与 $\tilde{P}_3^{(2)}$ 同向同速运动的和频光 \tilde{E}_3 ，可从恒在其身旁的 $\tilde{P}_3^{(2)}$ 中源源不断地汲取能量，而 $\tilde{P}_3^{(2)}$ 虽在被汲取后振幅又有所降低，但此时两个作为原料的混频光 $\tilde{E}_1 = E_{01} e^{i(k_1 z - \omega_1 t)}, \tilde{E}_2 = E_{02} e^{i(k_2 z - \omega_2 t)}$ 又会通过二阶非线性极化过程，源源不断

地转换为 $\tilde{\mathbf{P}}_3^{(2)}$ 以补充其因传输给 $\tilde{\mathbf{E}}_3$ 而损失的能量。所以总地来说，能量流动的方向和步骤为：起点 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2 \rightarrow$ 中转站 $\tilde{\mathbf{P}}_3^{(2)} \rightarrow$ 终点 $\tilde{\mathbf{E}}_3$ ；而从微观粒子的角度，此时 $\tilde{\mathbf{E}}_3$ 传播方向上每一个偶极子的振动频率与两两间的空间间隔的比值恰到好处，以使得每一个偶极子所释放出来的辐射恰好均彼此相干地加入了 $\tilde{\mathbf{E}}_3$ 中，以至于全体参与辐射的偶极子所辐射出的总能量应无损地正比于参与辐射的偶极子数目的平方。由于此时 $\tilde{\mathbf{E}}_3, \tilde{\mathbf{P}}_3^{(2)}$ 二者的波矢、频率相同，相位差固定，因此式(2.52)也称为完美相位匹配条件。

(2) 波矢失配时，和频波能流随波矢失配程度的增加以 $\text{sinc}^2(\cdot)$ 下降和振荡

当完美相位匹配条件即式(2.52)不满足时，考虑极化光 $\tilde{\mathbf{P}}_3^{(2)}$ 从 $z=0$ 开始向 $\tilde{\mathbf{E}}_3$ 转换 ($A_3(0)=0$)，并于 $z=L$ 结束转换过程，则对入射光为定振幅的式(2.42)积分得：

$$A_3(L) = A_3(L) - A_3(0) = \frac{i|\omega_3| \sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c} P_{03}^{(2)} \int_0^L e^{i\Delta k \cdot z} dz = \frac{i|\omega_3| \sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c} P_{03}^{(2)} \frac{e^{i\Delta k \cdot L} - 1}{i\Delta k} \quad (2.54)$$

设 $\tilde{\mathbf{E}}_3$ 从 $z=L$ 出射晶体后端面（或结束波矢匹配），则由于出射晶体（结束相位匹配）后振幅不再改变，则据(C205)得单色定振幅平面波 $\tilde{\mathbf{E}}_3$ 的能流密度时间平均值为：

$$\begin{aligned} \mathbf{I}_3 = \overline{\tilde{\mathbf{S}}_3} &= \frac{\mathbf{S}_3}{2} = \frac{1}{2} \mathcal{E}_3 \times \mathcal{H}_3 = 2\mathbf{E}_{03} \times \mathbf{H}_{03} = 2\sqrt{\frac{\epsilon_3^{(1)}}{\mu_3^{(1)}}} |\mathbf{E}_{03}|^2 \hat{\mathbf{k}}_3 = 2\epsilon_3^{(1)} \sqrt{\frac{1}{\mu_3^{(1)} \epsilon_3^{(1)}}} |A_3(L)|^2 \hat{\mathbf{k}}_3 = 2\epsilon_3^{(1)} v_3 \\ \cdot |A_3(L)|^2 \hat{\mathbf{k}}_3 &= 2\epsilon_0 \epsilon_{r3}^{(1)} \frac{c}{n_3} |A_3(L)|^2 \hat{\mathbf{k}}_3 = 2\epsilon_0 n_{3e}^2 \frac{c}{n_{3m} n_{3e}} |A_3(L)|^2 \hat{\mathbf{k}}_3 = \frac{2\epsilon_0 n_{3e} c}{n_{3m}} |A_3(L)|^2 \hat{\mathbf{k}}_3 \end{aligned} \quad (2.55)$$

同理对于两个参与混频的定振幅入射光 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ ，也有类似的：

$$\mathbf{I}_i = \frac{2\epsilon_0 n_{ie} c}{\sqrt{\mu_{ri}^{(1)}}} |E_{0i}|^2 \hat{\mathbf{k}}_i \quad (2.56)$$

将式(2.35)或(2.37)、式(2.54)代入式(2.55)得：

$$\begin{aligned} I_3 &= \frac{2\epsilon_0 n_{3e} c}{\sqrt{\mu_{r3}^{(1)}}} |A_3(L)|^2 = \frac{2\epsilon_0 n_{3e} c}{\sqrt{\mu_{r3}^{(1)}}} \frac{\omega_3^2 \mu_{r3}^{(1)}}{(2\epsilon_0 n_{3e} c)^2} \left| P_{03}^{(2)} \frac{e^{i\Delta k \cdot L} - 1}{\Delta k} \right|^2 = \frac{\omega_3^2 \sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c} |P_{03}^{(2)}|^2 \left| \frac{e^{i\Delta k \cdot L} - 1}{\Delta k} \right|^2 \\ &= \frac{\omega_3^2 \sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c} (4\epsilon_0 d_{eff})^2 \cdot |E_{01} E_{02}|^2 \left| \frac{e^{i\Delta k \cdot L} - 1}{\Delta k} \right|^2 = \frac{\omega_3^2 \sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c} (4\epsilon_0 d_{eff})^2 \cdot \frac{I_1 \sqrt{\mu_{r1}^{(1)}}}{2\epsilon_0 n_{1e} c} \frac{I_2 \sqrt{\mu_{r2}^{(1)}}}{2\epsilon_0 n_{2e} c} \left| \frac{e^{i\Delta k \cdot L} - 1}{\Delta k} \right|^2 \\ &= \frac{2d_{eff}^2 \omega_3^2 \sqrt{\mu_{r1}^{(1)} \mu_{r2}^{(1)} \mu_{r3}^{(1)}} I_1 I_2}{n_{1e} n_{2e} n_{3e} \epsilon_0 c^3} \cdot \left| \frac{e^{i\Delta k \cdot L} - 1}{\Delta k} \right|^2 \end{aligned} \quad (2.57)$$

其中

$$\left| \frac{e^{i\Delta k \cdot L} - 1}{\Delta k} \right|^2 = L^2 \left(\frac{e^{i\Delta k \cdot L} - 1}{\Delta k \cdot L} \right) \cdot \left(\frac{e^{i\Delta k \cdot L} - 1}{\Delta k \cdot L} \right)^* = L^2 \left(\frac{e^{i\Delta k \cdot L} - 1}{\Delta k \cdot L} \right) \cdot \left(\frac{e^{-i\Delta k \cdot L} - 1}{\Delta k \cdot L} \right) = L^2 \left[\frac{2 - 2 \cos(\Delta k \cdot L)}{(\Delta k \cdot L)^2} \right]$$

$$= 2L^2 \left[\frac{2\sin(\Delta k \cdot L/2)^2}{(\Delta k \cdot L)^2} \right] = L^2 \frac{\sin(\Delta k \cdot L/2)^2}{(\Delta k \cdot L/2)^2} = L^2 \text{sinc}^2\left(\frac{\Delta k \cdot L}{2}\right) \quad (2.58)$$

注意其中 sinc 函数是数学上而非数字信号处理中的 sinc 函数。于是式(2.57)变为：

$$I_3 = \frac{\sqrt{\mu_{r1}^{(1)} \mu_{r2}^{(1)} \mu_{r3}^{(1)}}}{n_{1e} n_{2e} n_{3e}} \frac{2d_{eff}^2 \omega_3^2 I_1 I_2}{\epsilon_0 c^3} \cdot L^2 \text{sinc}^2\left(\frac{\Delta k \cdot L}{2}\right) = I_{3\max} \text{sinc}^2\left(\frac{\Delta k \cdot L}{2}\right) \propto \text{sinc}^2\left(\frac{\Delta k \cdot L}{2}\right) \quad (2.59)$$

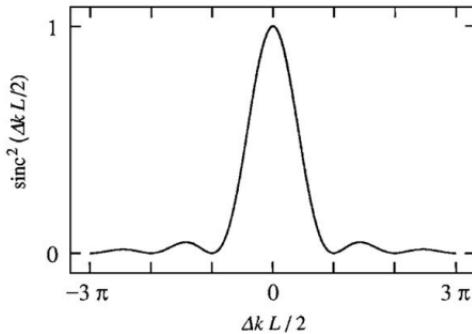


图 2.2 不同波矢失配程度下的和频波与完美相位匹配时的和频波的能量比

可见当 $|\Delta k \cdot L| = |\Delta k| \cdot L$ 增加时，三波混频过程的效率相对于随传播距离线性增长的完美相位匹配时的 $I_{3\max}$ 有所下降；当 $|\Delta k| \cdot L/2 > \pi$ 即 $L > 2\pi/|\Delta k|$ 后， \tilde{E}_3 的相位会与驱动它生成的非线性极化波源 $\tilde{P}_3^{(2)}$ 几乎完全脱离匹配。对式(2.42)取模，或者更好的是对式(2.59)开方并对 $z=L$ 求一阶导，给出随着 z 从 0 取到 $2\pi/|\Delta k|$ ， $A_3(z)$ 会先因 L 随 z 增大的增加速度大于 $\text{sinc}(\Delta k \cdot L/2)$ 随 z 增大的减小速度而先增大，而后因 L 的增大程度慢于 $\text{sinc}(\Delta k \cdot L/2)$ 的减小程度而减小，对应 $dA_3(z)/dz$ 先正后负，在 \tilde{E}_3 传播到 $0 \sim 2\pi/|\Delta k|$ 间的极值点 $d[L \cdot \text{sinc}(\Delta k \cdot L/2)]/dL = d[2/\Delta k \cdot \sin(\Delta k \cdot L/2)]/dL = \cos(\Delta k \cdot L/2) = 0$ ($\Delta k \neq 0$) 即

$$z_n = \frac{\pi n - \pi/2}{\Delta k/2} (n \in \mathbb{Z}) \quad (2.60)$$

处即介于完美相位匹配到完全脱离匹配的中点 $z_0 = \pi/|\Delta k|$ (一般 $\Delta k < 0$) 处时，能量会开始从 \tilde{E}_3 流回 $\tilde{P}_3^{(2)}$ 并返回到 \tilde{E}_1, \tilde{E}_2 中，因此定义其中 \tilde{E}_3 的净能量开始流出之时的长度 $L_{coh} := z_0 = \pi/|\Delta k|$ 为相干长度，于是式(2.59)还可写作：

$$I_3 = \frac{\sqrt{\mu_{r1}^{(1)} \mu_{r2}^{(1)} \mu_{r3}^{(1)}}}{n_{1e} n_{2e} n_{3e}} \frac{2d_{eff}^2 \omega_3^2 I_1 I_2}{\epsilon_0 c^3} \cdot L^2 \text{sinc}^2\left(\frac{\pi}{2} \frac{L}{L_{coh}}\right) = I_{3\max} \text{sinc}^2\left(\frac{\pi}{2} \frac{L}{L_{coh}}\right) \quad (2.61)$$

当 \tilde{E}_1, \tilde{E}_2 到和频光 \tilde{E}_3 的转换充分、转换效率较高时，和频光 \tilde{E}_3 从 \tilde{E}_1, \tilde{E}_2 中吸取的能量较多，且随着在晶体中的传播深度的增加而吸取能量增多，因此不仅会造成

$\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 能量和振幅的显著下降，并且此时 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 的振幅会随着非线性过程的进一步进行、随着 z 的增加而减小^①，以至于此时 $A_1(z), A_2(z)$ 是含空的，因此此时两个参与混频的入射光场的振幅不再是定振幅，导致不再有式(2.59)关系式成立。但即使在这种情况下，出射和频光场的极大值 $I_{3\max}$ 仍然是在完美相位匹配条件即(2.52)被满足时达到，极值点 $\Delta k = 0$ 没变。

注意即使波矢失配量 Δk 不为零，相应和频光的光强也只是相对于完美相位匹配的光强而言随传播距离的增长而下降得比较厉害，但若看其绝对能流密度或其绝对振幅，则是以 $2L_{coh}$ 为周期的正弦三角函数 $|\sin(\pi/2 \cdot L/L_{coh})|$ 的形式随传播距离振荡，若和频光在晶体内的传播距离、波矢失配距离恰好是相干长度 L_{coh} 的奇数倍，则和频光振幅仍可取得此情况下的最大值，其光强为 $(2/\Delta k)^2 = (2L_{coh}/\pi)^2$ 乘以式(2.61)的系数部分，在 Δk 较小但不为零的时候，仍然较大，只不过此时其振幅无法享受随 L 的线性增益，只能取分立的值 $L = (2m+1) \cdot L_{coh}$ 以保全自身的值不至于太小，但上限也就如此。

（3）实现相位匹配的两种方法

①. 通常情况难以实现(完美)相位匹配的原因

可以从以下过程看出完美相位匹配条件 $\Delta k = 0$ 在一般的正常色散非线性介质中很难被满足：在之前的克莱曼对称性成立的前提下，介质的色散可以被忽略，并且这导致介质是无损的。但在一般的介质中，即使介质是无损的，或多或少也是有色散的（即使色散可忽略），并且大多都符合正常色散规律，即一阶非线性系数张量的各分量、某方向上的等效标量光学折射率 $n(\omega)$ ，均随着晶体中光场频率 ω 的增加而增加。

对于两正频光的和频过程 $\omega_1 + \omega_2 = \omega_3$ ，总可以假设 $0 < \omega_1 \leq \omega_2 < \omega_3$ ，此时正常色散介质中有 $n_1 \leq n_2 < n_3$ 。另一方面，假设完美相位匹配条件 $\Delta k = k_1 + k_2 - k_3 = 0$ 可以满足，即下式可以成立

$$\frac{n_1\omega_1}{c} + \frac{n_2\omega_2}{c} = \frac{n_3\omega_3}{c} \quad (2.62)$$

那么 $n_1\omega_1 + n_2\omega_2 = n_3(\omega_1 + \omega_2)$ ，得到 $(n_1 - n_3)\omega_1 = (n_3 - n_2)\omega_2$ ，又因 $\omega_1 \leq \omega_2$ ，有

$n_1 - n_3 \geq n_3 - n_2$ ，即有：

^① 注意这不是介质有损的结果，而纯粹是介质的非线性系数较大以至非线性极化过程较大，并且同时因光子晶体内倒格矢较大程度满足波矢匹配条件，实现近似完美相位匹配，所导致的 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 到 $\tilde{\mathbf{P}}_3^{(2)}$ 、 $\tilde{\mathbf{P}}_3^{(2)}$ 到 $\tilde{\mathbf{E}}_3$ 的两个过程的转化效率均很高，才导致了 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 损失了过多能量并分配给了 $\tilde{\mathbf{E}}_3$ 。

$$n_1 + n_2 \geq 2n_3 \quad (2.63)$$

但在正常色散中有 $n_1 \leq n_2 < n_3$, 对应 $n_1 + n_2 \leq 2n_3$ 以及 $k_1 + k_2 \leq k_3$, 所以正常色散中不可能实现完美相位匹配; 从相速度的角度也可理解这一点^①。

②. 利用反常色散实现相位匹配

所有的介质都是色散介质的同时, 几乎所有的介质也都是反常色散介质, 只是大多数晶体是透明且没有颜色的, 以至于这些介质在远离其固有频率的可见光区内都表现出正常色散, 晶体该区域频率的光没有选择性吸收, 对各频率电磁波的吸收程度都基本均匀一致; 但实际上这些介质在不可见光区大多都有反常色散区, 这个区域在介质的固有频率附近, 对应晶体对与之固有频率相匹配的电磁波的选择性吸收。

因此若将 $\omega_1 + \omega_2 = \omega_3$ 从远离晶体固有频率的甚低频区, 增加频率至接近晶体固有频率时, 在步入反常色散区间后, n_3 会随着 ω_3 的增加而下降, 而 n_1, n_2 则由于还未到达反常色散区间或超过得不多, 而继续增加或下降得并不很明显, 以至 $n_1 + n_2 \leq 2n_3$ 左侧的值保持不减或小幅下降, 而右侧的值却大幅下降, 以至参与混频的三波的频率所对应的折射率最终可以满足(2.63)的关系式, 以至完美相位匹配变得能够实现。

但当 ω_3 增大至反常色散区间时, 介质不仅对 ω_3 有强烈的吸收导致生成的 ω_3 振幅大幅降低, 并且由于此时 ω_1, ω_2 中至少有一个的值是 $> \omega_3/2$ 的, 那么这两束入射光中至少有一束也会开始被介质吸收而导致振幅降低, 以至整个和频过程的转换效率大幅下降, 并且由于晶体的强烈吸收导致的热效应等还有损晶体结构及其性能, 更有甚者会导致克莱曼对称性的失效, 甚至连无吸收介质中的非线性系数的对称性都在此情况下失效了, 于是上述描述非线性过程的耦合波方程等均完全失效, 以至上述现有理论无法精确地描述利用反常色散实现的相位匹配过程和结果。所以反常色散虽理论上能实现相位匹配, 但实际上很少有应用。

③. 实现相位匹配的最常用手段: 非线性各向异性晶体的双折射

^① 根据 $n_1(\omega_2 - \omega_1) \leq n_2(\omega_2 - \omega_1)$, 有 $n_1\omega_2 + n_2\omega_1 \leq n_1\omega_1 + n_2\omega_2$, 两边同除以 c 即有

$$\frac{\omega_2}{\omega_1}k_1 + \frac{\omega_1}{\omega_2}k_2 \leq k_1 + k_2, \text{ 得到 } 2\sqrt{k_1 k_2} \leq \frac{\omega_2}{\omega_1}k_1 + \frac{\omega_1}{\omega_2}k_2 \leq k_1 + k_2 \leq k_3, \text{ 于是 } (\sqrt{\frac{\omega_2}{\omega_1}k_1} + \sqrt{\frac{\omega_1}{\omega_2}k_2})^2 =$$

$$\frac{\omega_2}{\omega_1}k_1 + \frac{\omega_1}{\omega_2}k_2 + 2\sqrt{k_1 k_2} \leq (\sqrt{k_1} + \sqrt{k_2})^2 = k_1 + k_2 + 2\sqrt{k_1 k_2} \leq k_1 + k_2 + \frac{\omega_2}{\omega_1}k_1 + \frac{\omega_1}{\omega_2}k_2 \leq 2k_3, \text{ 最后}$$

一个不等式两边除以 ω_3 即得相速度关系 $v_1 + v_2 \leq 2v_3$, 其与 $n_1 + n_2 \leq 2n_3$ 、 $k_1 + k_2 \leq k_3$ 类似。

属于三方晶系的铌酸锂晶体不仅是非线性电介质，也是各向异性晶体，即其二阶非线性系数不为零的同时，一阶非线性系数也不为零。这样的晶体便可能利用双折射实现相位匹配，其原理是通过材料的双折射补偿介质的正常色散，以减小正常色散所带来的波矢失配程度。因此这种方法可以在远离材料的固有频率即共振吸收区下完成相位匹配，以至克莱曼对称性仍有效，则之前所导出的公式仍能较好地描述实验现象。

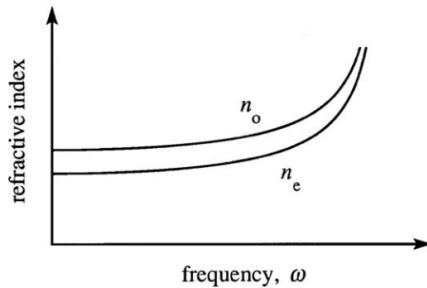


图 2.3 负单轴晶体某波矢或极化方向上 o 光 e 光的折射率各自的正常色散关系

由于介质的二阶非线性系数与一阶非线性系数的取值是无关的，即介质是否是非线性与其是否是各向异性无关，比如自然界存在是非线性晶体却又同时是各向同性晶体的介质：属于立方晶系的晶体就是光学各向同性的，即使其同时也是非线性的，则由于没有双折射现象便无法利用双折射实现相位匹配。

表 2.2 对不同晶系的一阶非线性系数的分类

System	Linear Optical Classification
Triclinic, monoclinic, orthorhombic	Biaxial
Trigonal, tetragonal, hexagonal	Uniaxial
Cubic	Isotropic

为了实现相位匹配，关系式(2.63)需要首先得到满足，因此需要想尽办法降低 n_3 和增大 n_1, n_2 。而各向异性介质提供了天然的随光波的传播方向或极化方向改变的折射率分布：以单轴晶体为例，其折射率椭球和 o 光 e 光的折射率曲面均可描述其折射率分布。前者是描述晶体中不同极化方向所对应的等效标量光学折射率，对于铌酸锂这一负单轴晶体而言， $\varepsilon_{n\perp}^{(1)} > \varepsilon_{n\parallel}^{(1)}$ ，因此其折射率椭球所对应的旋转椭球面的短轴即旋转轴即光轴，那么为了使得 $n(\omega_3)$ 更小的同时 $n(\omega_1), n(\omega_2)$ 较大， ω_3 的极化方向应尽量朝着平行于光轴的方向，同时 ω_1, ω_2 的极化方向应尽量垂直于光轴方向；

若考虑后者即 o 光、e 光的折射率曲面，则由于二者均相切于垂直于光轴的方向，在以铌酸锂晶体为代表的负单轴晶体中，从图 2.3（其来源需察看附录 C 中对图 C2 的推导）可见，二者相切于 e 光折射率曲面的长轴，此时在任何波矢方向上，e 光的

等效标量光学折射率 $n_{ne}/\sqrt{\mu_{nr}^{(1)}} = n_{ne} < n_{neo} = n_{no}/\sqrt{\mu_{nr}^{(1)}}$ ，因此从这个角度， ω_3 必须是 e 光，并且其波矢方向应尽量垂直于光轴（这与其极化方向最好平行于光轴并不矛盾），同时 ω_1, ω_2 最好均是 o 光，这是最容易实现相位匹配的组合，此时若 ω_1, ω_2 的入射面均与主截面重合，则 ω_3 的极化方向与 ω_1, ω_2 的极化方向互相垂直、 ω_1, ω_2 的极化方向之间的夹角等于二者在晶体内的波矢夹角。

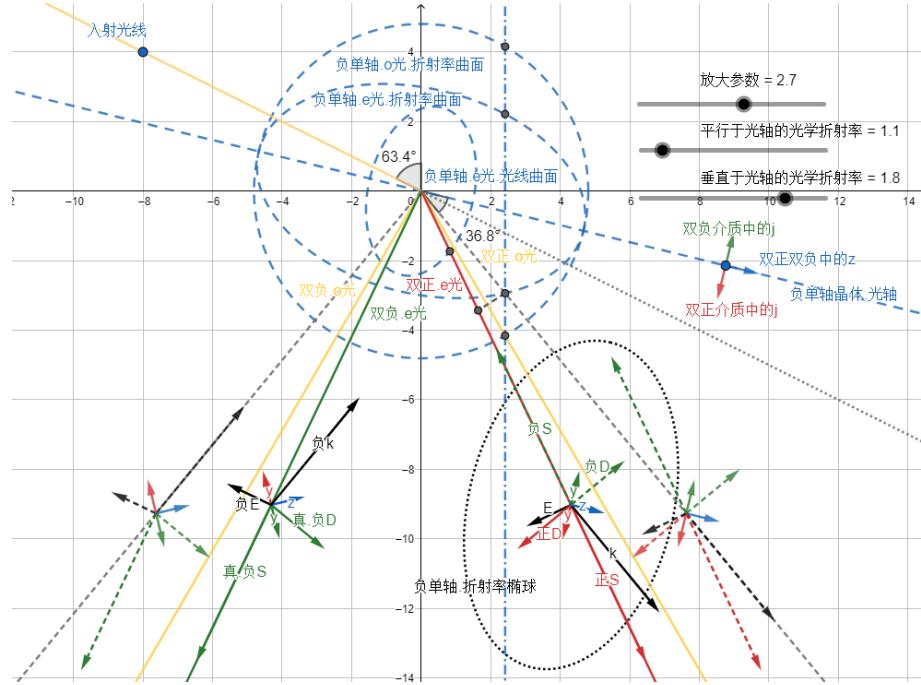


图 2.4 单一入射光下，双正和双负单轴晶体中的入射光、正常和反常折射时的 o 光 e 光

另外，除了 ω_1, ω_2 为双 o 光外，也可以一个是 o 光一个是 e 光（不能是两个 e 光，否则三光均遵循 e 光的正常色散曲线将导致式(2.63)不成立），在 $0 < \omega_1 \leq \omega_2$ 的假设下， ω_1 为 e 光而 ω_2 保持 o 光的组合，会比 ω_1, ω_2 为双 o 光的组合，对于实现相位匹配要稍微难一点，但比 ω_2 为 e 光、 ω_1 为 o 光的组合更容易实现相位匹配，即为满足相位匹配，并不如后者那么需要 n_{ne}, n_{no} 相对差值 $|n_{ne} - n_{no}|/n_{no}$ 得较大。同理，正单轴晶体中难易程度排序的得来过程类似，因此有：

表 2.3 单轴晶体的相位匹配方法

已知	ω_1	+	ω_2	=	ω_3	频率匹配
实现	$n_1\omega_1$	+	$n_2\omega_2$	=	$n_3\omega_3$	波矢匹配
方法	n_1		n_2		n_3	组合类型
负单轴	o		o		e	Type I

e	o	e	Type II	中
o	e	e	Type III	难
e	e	o	Type I	易
正单轴	o	e	Type II	中
e	o	o	Type III	难

具体的通过双折射实现相位匹配的办法，可分为角度调节法和温度调节法。

1. 角度调节法

该方法可给出完美相位匹配时，晶体光轴与折射倍频 e 光波矢间的夹角。据式(C107)，将其左右两侧同除以 $\mu_{ne}^{(1)}$ ，得各向同性的单轴晶体中的 e 光的折射率曲面：

$$\frac{1}{n_{ne}^2(\theta_{ne})} = \frac{\sin^2 \theta_{ne}}{n_{n//}^2} + \frac{\cos^2 \theta_{ne}}{n_{n\perp}^2} \quad (2.64)$$

为简便起见，考虑铌酸锂负单轴晶体中 I 型二次谐波/倍频波的生成过程。对于二次谐波即 $\omega_1 = \omega_2 = \omega$ 、 $\omega_3 = \omega_1 + \omega_2 = 2\omega$ 的和频波的生成过程，据表 2.3，波矢匹配写为 $2n(\omega) \cdot \omega = n(2\omega) \cdot (2\omega)$ ，即 $n(\omega) = n(2\omega)$ ，又由于是 I 型的，则有 $n_o(\omega) = n_e(2\omega)$ 即：

$$n_e(2\omega, \theta_e) = n_o(\omega) \quad (2.65)$$

同时又因 o 光折射率曲面为 $n_o(\omega) = n_\perp(\omega)$ ，将其与式(2.65)一同代入式(2.64)得

$$\begin{aligned} \frac{1}{n_\perp^2(\omega)} &= \frac{1}{n_o^2(\omega)} = \frac{1}{n_e^2(2\omega, \theta_e)} = \frac{\sin^2 \theta_e(2\omega)}{n_{n//}^2(2\omega)} + \frac{\cos^2 \theta_e(2\omega)}{n_{n\perp}^2(2\omega)} \\ &= \frac{\sin^2 \theta_e(2\omega)}{n_{n//}^2(2\omega)} + \frac{1 - \sin^2 \theta_e(2\omega)}{n_{n\perp}^2(2\omega)} \end{aligned} \quad (2.66)$$

即有

$$\sin^2 \theta_e(2\omega) = \frac{\frac{1}{n_\perp^2(\omega)} - \frac{1}{n_\perp^2(2\omega)}}{\frac{1}{n_{n//}^2(2\omega)} - \frac{1}{n_{n\perp}^2(2\omega)}} \quad (2.67)$$

该式便给出了达成完美相位匹配时晶体晶轴与晶体内折射光中倍频 e 光波矢间的夹角 $\theta_e = \langle \mathbf{k}_e, \mathbf{k} \rangle$ ，即展示了晶体该如何旋转以实现相位匹配。

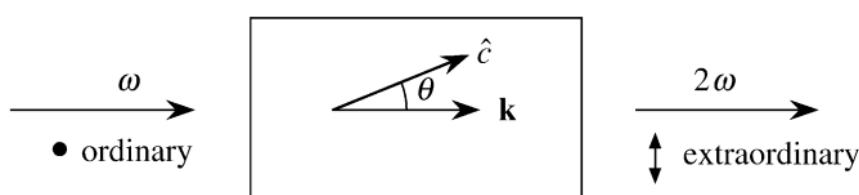


图 2.5 负单轴晶体中二次谐波生成过程中为实现相位匹配的角度调节几何示意图

但由于有些材料正常色散曲线在低频近似直线的部分，斜率过大，或双折射效应太小，会导致方程(2.67)右侧大于单位 1，以至角度 θ_e 无实数解，此时方程的解 θ_e 便没有物理意义。因此单纯地旋转晶体或改变基频光入射方向，对于某些晶体可能仍无法实现相位匹配。

2. 温度调节法

角度调节法还有一个缺点，据式(C128)， e 光的离散角 $\alpha_e = \langle \tilde{\mathbf{S}}_e, \mathbf{k}_e \rangle$ 在 $\theta_e = \langle \mathbf{k}_e, \mathbf{k} \rangle$ 不为零也不为 90° 时，恒不为零，如图 2.3 所示。意味着通常情况下，身为 e 光的倍频光的能流密度矢量与其波矢方向不同，而在讨论的 I 型波矢匹配中，基频光是 o 光，其能量密度矢量与其波矢量方向相同 $\alpha_o = \langle \tilde{\mathbf{S}}_o, \mathbf{k}_o \rangle = 0$ ；又因基频光极化所得的非线性倍频极化光的波矢量等于基频光的波矢量的两倍，而在完美相位匹配条件下，二次谐波也与作为能量传递媒介的倍频极化光的波矢相同，因此此时在晶体中，均作为折射光的基频光、倍频光、倍频极化光三者的波矢均同向，即有 $\langle \tilde{\mathbf{S}}_o, \mathbf{k}_e \rangle = \langle \tilde{\mathbf{S}}_o, \mathbf{k}_o \rangle$ ，于是：

$$\langle \tilde{\mathbf{S}}_e, \tilde{\mathbf{S}}_o \rangle = \langle \tilde{\mathbf{S}}_e, \mathbf{k}_e \rangle - \langle \tilde{\mathbf{S}}_o, \mathbf{k}_e \rangle = \alpha_e - \langle \tilde{\mathbf{S}}_o, \mathbf{k}_o \rangle = \alpha_e - \alpha_o = \alpha_e \neq 0 (\theta_e \neq 0, 90^\circ) \quad (2.68)$$

此时 o 光倍频极化光与 e 光倍频光的能流方向从生成了 e 光那一刻起，就迅速从 e 光生成点开始，随着继续往前传播而彼此分离。而能流密度的越分越开意味着行波电磁场的交叠程度越来越小，以至于降低了依赖叠加场来随传播距离以传输能量的任何非线性混频过程的转换效率，并从分开之后倍频光场振幅不再随传播距离增加，而是随着振幅在最大振幅基础上加上 $|\sin(\pi/2 \cdot L/L_{coh})|$ 形式的增量后上下振荡，继续单独走完晶体中最后一段距离并出射晶体。在此过程中二次谐波在出射晶体前就混频完成，混频从始到终的有效长度 L 很短，没有利用完晶体的剩余长度和全长，也没有与基波充分交叠并从中汲取能量，以至最终出射晶体的二次谐波的场振幅非常小。这种效应被称为走离效应。

但对于有些晶体如铌酸锂晶体而言，其双折射程度强烈依赖于温度，因此可以通过保持 $\theta_e = \langle \mathbf{k}_e, \mathbf{k} \rangle = 90^\circ$ 不变，并改变晶体温度来实现混频过程中的完美相位匹配。

2.2 一维非线性光子晶体中的准相位匹配机制

有两种比较典型的情况无法使用双折射效应实现相位匹配。一种是当材料在所考虑的波长或频率范围内有较大非线性系数或效应，但没有双折射效应，或双折射效应

太弱而无法补偿一阶非线性系数的正常色散时。比如高频时 $|n_e(2\omega) - n_o(2\omega)|/n_o(2\omega)$ 急剧下降， $n_e(2\omega)$ 与 $n_o(2\omega)$ 区别不大，以至于即使对倍频光、和频光选用 e 光，双折射效应对折射率系数的正常色散的补偿作用也微不足道，波矢最终还是会失配。

还有一种通过双折射实现相位匹配的办法失效的情况是，有时候针对特定的应用（如接下来的利用铌酸锂晶体所进行的非线性全息），会想办法使用比非对角二阶非线性系数大很多的对角系数 d_{33} ，但要想让 d_{33} 出现在 d_{eff} 中，必须让参与混频的两条入射光 \tilde{E}_1, \tilde{E}_2 的偏振方向相同且均为 e 光，那么根据表 2.3，便已没有符合条件的双折射相位匹配类型了。当然，之后会给出此时 $\tilde{P}_3^{(2)}, \tilde{E}_3$ 的振动方向也与 \tilde{E}_1, \tilde{E}_2 振动方向大致相同，而由于傍轴近似条件下 $\tilde{E}_1, \tilde{E}_2, \tilde{P}_3^{(2)}, \tilde{E}_3$ 四者的波矢大致同向，只要光轴方向不与四者的波矢近似同向，则四者的主平面也大致重合，因此此时即使有双折射效应、且即使双折射效应还很强，也没法利用上双折射效应来实现相位匹配了，否则参与混频的三光的极化和偏振方向就不大致全同了，所以此时四者必须同时全为 o 光或 e 光，从这一点也可看出双折射相位匹配技术失效。

此时还有一个好处，那就是当 \tilde{E}_3 与 \tilde{E}_1, \tilde{E}_2 的波矢与主平面都大致重合时，无论三者是 o 光还是 e 光，因 \tilde{E}_3 与 \tilde{E}_1, \tilde{E}_2 的波矢与光轴的夹角近似相同，则 \tilde{E}_3 与 \tilde{E}_1, \tilde{E}_2 的离散角及其所确定的能流密度相对于波矢的方向，以及因主平面相同而能流的绝对方向，也都大致相同。这样 \tilde{E}_3 与 \tilde{E}_1, \tilde{E}_2 的交叠程度在晶体内短程有效相互作用距离内不会下降太多，保证了和频和倍频过程的顺利进行。

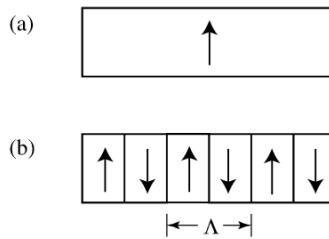


图 2.6 自然均匀单向自发极化的铁电晶体(a)和人工周期性极化的铁电材料(b)

当普通的相位匹配技术无法实施时，准相位匹配技术就大显神通了：利用人工工艺将一维铁电材料沿 z 轴方向以 $\Lambda/2 = L_{coh} = \pi/|\Delta k|$ 为厚度将其切成一片片薄层，并以 $\Lambda = 2L_{coh} = 2\pi/|\Delta k|$ 为周期将相邻两片薄层的自发极化方向即晶轴(相对于波矢的)正向翻转 180° 排列（图 2.6），而晶轴方向的翻转意味着原自然单向均匀极化晶体的 d_{eff} 反号，便可因此获得受矩形波函数调制的随 z 变化的二阶非线性极化率系数

$$d(z) = d_{eff} \cdot \text{sgn} \left[\cos \left(\frac{2\pi}{\Lambda} z \right) \right] = d_{eff} \cdot \left\{ \sum_{n=-\infty}^{+\infty} 2 \text{rect} \left[\frac{2(z-n\Lambda)}{\Lambda} \right] - 1 \right\}$$

$$= d_{\text{eff}} \cdot \left\{ 2 \sum_{n=-\infty}^{+\infty} \text{rect}\left(\frac{2z - 2n\Lambda}{\Lambda}\right) - 1 \right\} \quad (2.69)$$

2.2.1 准相位匹配是除完美相位匹配方法外提高转换效率的较好方法

2.2.1.1 二阶非线性系数的空间周期调制函数可展成不同倒格矢的倒格波的线性叠加

根据图 2.2 和式(2.69)，每当和频光或倍频光传播半个周期即一个薄膜厚度，振幅因波矢失配而快要开始衰减之时 ($L = (2m+1) \cdot z_0 = (2m+1) \cdot L_{\text{coh}}$)，立刻使其步入下一个自发极化方向相反薄膜中，随着非线性系数的反号，能量的绝对流向也反向，以至当同一地点处倍频光的相位超前于极化光相位 $\pi/2$ 内时，能量朝 k_3 正向流动，而当倍频光相位滞后于极化光 $\pi/2$ 内时，能量逆 k_3 反向流动，这样就保证了能量的相对流向始终从基频光流向倍频光，维持二次谐波振幅随传播距离而持续增长，如下图所示：

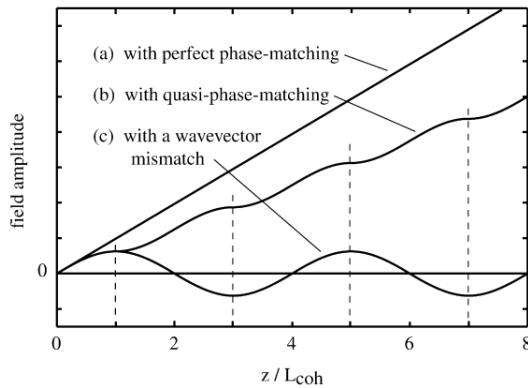


图 2.7 三种相位匹配条件下生成的和/差/倍频波振幅随传播距离的变化对比

接下来将从数学上描绘和证明这一点。由于在这样的人工材料中二阶非线性极化率系数从 d_{eff} 变为了式(2.69)的 $d(z)$ ，因此原耦合波方程等一系列方程中的 d_{eff} 也需要修正为 $d(z)$ 。现将 $d(z)$ 展为复数形式傅里叶级数：

$$d(z) = d_{\text{eff}} \cdot \sum_{m=-\infty}^{+\infty} C_m e^{iG_m z} \quad (2.70)$$

其中级数部分来源于将调制函数 $\text{sgn}[\cos(2\pi/\Lambda \cdot z)]$ 展为复数形式的傅里叶级数：

$$\text{sgn}\left[\cos\left(\frac{2\pi}{\Lambda} z\right)\right] = \sum_{m=-\infty}^{+\infty} C_m e^{\frac{i m \pi z}{\Lambda/2}} = \sum_{m=-\infty}^{+\infty} C_m e^{\frac{i 2\pi m}{\Lambda} z} = \sum_{m=-\infty}^{+\infty} C_m e^{iG_m z} \quad (2.71)$$

其中 $k_m = 2\pi m / \Lambda$ 为 $d(z)$ 的第 m 级傅里叶分量的光栅矢量，而 $2\pi/\Lambda$ 等效于拥有简立方、面心立方、体心立方正格子晶胞（三者的“外壳”8 顶点均构成简立方）或长方体正格子晶胞点阵构型的一维光子晶体的倒格子的平行于光轴的基矢的大小。因此 $k_m = 2\pi m / \Lambda$ 可以看做一维光子晶体的一维倒易点阵中的某两点间的倒格矢，但不能将其比作光子晶体的格波波矢，因为格波波矢的分母是一维原子链长度，而这里光子晶

体倒格矢的分母不是一维光子晶体厚度，而是二阶非线性系数的调制函数在一个周期内的长度，前者比后者多了 N 倍，N 为正空间中的原子数或调制周期数。在这样的理解下，式(2.70)便有了物理意义，即其可以被认为是将二阶非线性系数展开为了以各级倒格矢为波矢量的各级倒格波（非正空间的格波，也非倒空间的格波即简正模）的线性叠加。

而根据式(2.69)，可通过傅里叶变换得傅里叶级数各级系数：

$$C_m = \mathcal{F} \left\{ \operatorname{sgn} \left[\cos \left(\frac{2\pi}{\Lambda} z \right) \right] \right\} = \frac{1}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} \left(2 \sum_{n=-\infty}^{+\infty} \operatorname{rect} \left(\frac{2z - 2n\Lambda}{\Lambda} \right) - 1 \right) e^{-iG_m z} dz \quad (2.72)$$

其中先考虑

$$\mathcal{F} \left\{ \sum_{n=-\infty}^{+\infty} \operatorname{rect} \left(\frac{2z - 2n\Lambda}{\Lambda} \right) \right\} = \sum_{n=-\infty}^{+\infty} \mathcal{F} \left\{ \operatorname{rect} \left(\frac{2z - 2n\Lambda}{\Lambda} \right) \right\} \quad (2.73)$$

其中

$$\begin{aligned} \mathcal{F} \left\{ \operatorname{rect} \left(\frac{2z - 2n\Lambda}{\Lambda} \right) \right\} &= \frac{1}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} \operatorname{rect} \left(\frac{2z - 2n\Lambda}{\Lambda} \right) e^{-iG_m z} dz \\ &= e^{-iG_m n\Lambda} \cdot \frac{1}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} \operatorname{rect} \left(\frac{2z - 2n\Lambda}{\Lambda} \right) e^{-iG_m (z-n\Lambda)} d(z-n\Lambda) \\ &= e^{-iG_m n\Lambda} \cdot \frac{1}{\Lambda} \int_{-\Lambda/2-n\Lambda}^{\Lambda/2-n\Lambda} \operatorname{rect} \left(\frac{2z}{\Lambda} \right) e^{-iG_m z} dz = e^{-iG_m n\Lambda} \cdot \frac{1}{2} \int_{-\Lambda/2-n\Lambda}^{\Lambda/2-n\Lambda} \operatorname{rect} \left(\frac{2z}{\Lambda} \right) e^{-i\pi m \frac{2z}{\Lambda}} d \left(\frac{2z}{\Lambda} \right) \\ &= e^{-iG_m n\Lambda} \cdot \frac{1}{2} \int_{-1-2n}^{1-2n} \operatorname{rect}(z) e^{-i\pi m z} dz \end{aligned} \quad (2.74)$$

当 $n \neq 0$ 时，积分区间上的 $\operatorname{rect}(z) = 0$ ，积分值为零，则式(2.73)只剩 $n=0$ 一项，以至原式(2.72)变为：

$$\begin{aligned} C_m &= 2 \mathcal{F} \left\{ \operatorname{rect} \left(\frac{2z}{\Lambda} \right) \right\} - \mathcal{F} \{ 1 \} = \int_{-1}^1 \operatorname{rect}(z) e^{-i\pi m z} dz - \frac{1}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} e^{-iG_m z} dz \\ &= \int_{-1/2}^{1/2} e^{-i\pi m z} dz - \frac{1}{2} \int_{-1}^1 e^{-i\pi m z} dz = \operatorname{sinc} \left(\frac{\pi m}{2} \right) - \operatorname{sinc}(\pi m) \end{aligned} \quad (2.75)$$

它的离散取值分布图像为：

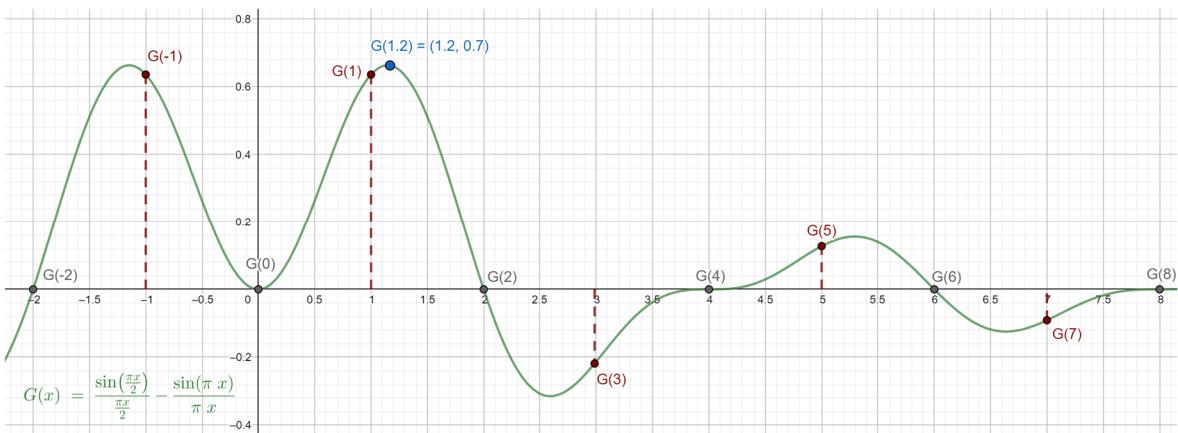


图 2.8 C_m 的非零值 C_{2k+1} ($k \in \mathbb{Z}$) 中 $C_{\pm 1}$ 最大

2.2.1.2 选择振幅较大的相应级次倒格波的波矢进行波矢匹配并以此设计调制周期可提高出射和频光场振幅

由于将 $d(z)$ 或其调制函数展开成傅里叶级数而不是傅里叶积分，因此 m 是离散取值而非连续取值的，因此即使 C_m 的极大值位置在 $m = \pm 1.2$ 左右，对于只取整数的 m ， C_m 的极大值仍然为 $C_{\pm 1}$ ，因此在 $d(z)$ 的级数展开式中，第 ± 1 级倒格矢所对应的倒格波的振幅是最大的，能量也是最高的，占据了 $d(z)$ 的绝大部分。因此即使任何调制函数的傅里叶级数展开式中任何一个倒格波的波矢都可以被用于补偿波矢失配，并以此来设计调制函数的调制周期，但由于式(2.71)形式的调制函数下的级数展开式的第 ± 1 级倒格波矢的振幅 $C_{\pm 1}$ 较高，为了使得所设计的 $d(z)$ 中补偿波矢失配的那个倒格波振幅最大，以提高生成的和频波振幅（采用其他级次的倒格波也可以，波矢失配总可以被满足，只不过调制周期需要相应地改变；但问题是其他级次的倒格波振幅不够高，导致出射光振幅不高，转换效率比使用该周期性结构下最有效的 1 级倒格波更低），必须采用 $d(z)$ 展开式中的第 ± 1 级倒格波来补偿波矢失配。

在采用了 ± 1 级倒格矢补偿波矢失配并以此设计了调制周期后，其他倒格波的波矢便不能准确补偿波矢失配，以至于即使波矢失配量中加上了这些倒格波波矢后，其相应的耦合波方程仍有不同程度的波矢失配，则这些倒格波所对应生成的和频光振幅 $A_3(z)$ 会随 z 的增加而在零附近以 $|\sin(\Delta k \cdot L/2)|$ 即 $|\sin(\pi/2 \cdot L/L_{coh})|$ 形式振荡，但最终出射晶体时仍会可能有小部分贡献。因此从一定的传播距离来看， ± 1 级倒格波之外的倒格波所生成的和频波对最终出射的和频波也有贡献，但相比 ± 1 级贡献比例不大且各异。而对于 ± 1 级也只取这两级中的某一级倒格波，因为级数 m 相反导致二者的波矢方向是反向的，所以只可能有一个对波矢失配起主要的补偿作用。于是式(2.70)变为：

$$d(z) \approx d_{eff} \cdot C_m e^{iG_m z} \quad (2.76)$$

将其代入三个耦合波方程(2.43)、(2.50-2.51)并替换掉其中的 d_{eff} ，得

$$\left\{ \begin{array}{l} \frac{dA_3(z)}{dz} = \frac{2i|\omega_3|\sqrt{\mu_{r3}^{(1)}}}{n_{3e}c} \left(d_{eff} \cdot C_m e^{iG_m z} \right) \cdot A_1(z) A_2(z) e^{i\Delta k \cdot z} = \frac{2i|\omega_3|\sqrt{\mu_{r3}^{(1)}} d_Q}{n_{3e}c} A_1(z) A_2(z) e^{i\Delta k_Q \cdot z} \\ \frac{dA_1(z)}{dz} = \frac{2i|\omega_1|\sqrt{\mu_{r1}^{(1)}} d_Q}{n_{1e}c} A_3(z) A_2^*(z) e^{iG_m z} e^{-i\Delta k \cdot z} = \frac{2i|\omega_1|\sqrt{\mu_{r1}^{(1)}} d_Q}{n_{1e}c} A_3(z) A_2^*(z) e^{-i(\Delta k_Q - 2G_m) \cdot z} \\ \frac{dA_2(z)}{dz} = \frac{2i|\omega_2|\sqrt{\mu_{r2}^{(1)}} d_Q}{n_{2e}c} A_3(z) A_1^*(z) e^{iG_m z} e^{-i\Delta k \cdot z} = \frac{2i|\omega_2|\sqrt{\mu_{r2}^{(1)}} d_Q}{n_{2e}c} A_3(z) A_1^*(z) e^{-i(\Delta k_Q - 2G_m) \cdot z} \end{array} \right. \quad (2.77)$$

以及参与混频的两入射光为定振幅时的：

$$I_3 = \frac{\sqrt{\mu_{r1}^{(1)} \mu_{r2}^{(1)} \mu_{r3}^{(1)}}}{n_{1e} n_{2e} n_{3e}} \frac{2\omega_3^2}{\epsilon_0 c^3} d_Q^2 \cdot I_1 I_2 \cdot L^2 \text{sinc}^2\left(\frac{\Delta k_Q \cdot L}{2}\right) \quad (2.78)$$

以及据式(2.56)有：

$$\begin{aligned} |E_{03}| &= \sqrt{\frac{\sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c}} I_3 = \sqrt{\frac{\sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c}} \frac{\sqrt{\mu_{r1}^{(1)} \mu_{r2}^{(1)} \mu_{r3}^{(1)}}}{n_{1e} n_{2e} n_{3e}} \frac{2\omega_3^2}{\epsilon_0 c^3} d_Q^2 \left(\frac{2\epsilon_0 n_{1e} c}{\sqrt{\mu_{r1}^{(1)}}} |E_{01}|^2 \frac{2\epsilon_0 n_{2e} c}{\sqrt{\mu_{r2}^{(1)}}} |E_{02}|^2 \right) \cdot L^2 \text{sinc}^2\left(\frac{\Delta k_Q \cdot L}{2}\right) \\ |E_{03}| &= |A_3(L)| = \frac{2\sqrt{\mu_{r3}^{(1)}} |\omega_3|}{n_{3e} c} d_Q \cdot |E_{01}| |E_{02}| \cdot L \cdot \text{sinc}\left(\frac{\Delta k_Q \cdot L}{2}\right) \propto d_Q \cdot |E_{01}| |E_{02}| \cdot L \cdot \left| \text{sinc}\left(\frac{\Delta k_Q \cdot L}{2}\right) \right| \end{aligned} \quad (2.79)$$

其中

$$d_Q = d_{eff} \cdot C_m \quad (2.80)$$

$$\Delta k_Q = \Delta k + G_m \quad (2.81)$$

而其中对于正常色散介质而言，波矢失配量 $\Delta k = k_1 + k_2 - k_3 < 0$ ，因此要想得到补偿后的波矢失配量 $\Delta k_Q = \Delta k + G_m = 0$ ，需有 $G_m = -\Delta k > 0$ ，因此应取正级次的倒格波；又由于想利用拥有最大振幅 $C_{\pm 1}$ 的 ± 1 级倒格波，则需取其中的+1 级倒格波，此时 $d_Q = d_{eff} \cdot C_m$ 才最大， $dA_3(z)/dz$ 及最后生成的 $A_3(L) \propto d_{eff} C_m \cdot |E_{01}| |E_{02}| \cdot L \cdot \text{sinc}(\Delta k_Q \cdot L/2)$ （设两入射光定振幅）才最大，对应图 2.7 中(b)曲线的斜率最高，即使因 $C_m < 1$ 而斜率没有(a)高。此时才有 $G_m = G_1 = 2\pi/\Lambda > 0$ ，同时 $C_m = C_1 = 2/\pi$ 。于是上述两式变为：

$$d_Q = d_{eff} \cdot C_1 = \frac{2}{\pi} d_{eff} \quad (2.82)$$

$$\Delta k_Q = \Delta k + \frac{2\pi}{\Lambda} \quad (2.83)$$

因此准相位匹配达成的条件为 $\Delta k_Q = \Delta k + 2\pi/\Lambda = 0$ ，即：

$$\Lambda = \frac{2\pi}{-\Delta k} = 2L_{coh} = \frac{2\pi}{k_3 - k_1 - k_2} \quad (2.84)$$

2.2.1.3 以周期性调制的二阶非线性系数中不同所选级次的倒格矢补上波矢失配量为目标地设计一维光子晶体可获得不同出射和频光振幅

该式子便提供了对于给定的三束混频光的波矢或波矢失配量，为弥补和消除波矢失配以达波矢匹配和相位匹配，一维光子晶体的二阶非线性系数的调制周期的计算和设计方法。在相同的调制函数和调制方法如(2.71)下，也可以采取不同级次的倒格波的倒格矢来进行波矢匹配，此时 $d_Q = d_{eff} \cdot C_m$ 会因所选倒格波的振幅 C_m 的变化而变化，以致影响 $dA_3(z)/dz$ 并最终影响 $A_3(z)$ 的大小，同时调制周期 $\Lambda = 2\pi m/|\Delta k|$ 也会有所变化。

一般地，参与混频的两束入射光的频率和波矢的方向和大小(即波长)是给定的，那么若只考虑二者分别在负单轴晶体中的 e 光成分或 o 光成分，则二者在晶体中的频率和波矢也便确定了（见图 2.4），而由于待合成的 ω_3 光的频率 $\omega_3 = \omega_1 + \omega_2$ 也确定了，则 ω_3 光的波矢大小，随考虑其为 e 光或 o 光，也通过关系式 $k_3 = |\omega_3|/v_3 = |\omega_3|n_3/v_3 = |\omega_3|\sqrt{\mu_{r3}^{(1)}}n_{3ee}/c$ 或 $|\omega_3|\sqrt{\mu_{r3}^{(1)}}n_{3eo}/c$ 确定了，但问题就来了，待合成的和频光 ω_3 的波矢方向没有确定！这一点非常关键，它直接暗示了三维情况下，在波矢失配满足时，将可能不仅有一条倍频光出射，而是所有波矢失配量被三维调制后的二阶非线性系数展开所得的各级倒格波的倒格矢所补偿上了的 ω_3 光，即所有以极化光的波矢终点为起点的倒格矢的终点，若落在以极化光的波矢起点为中心所作的旋转椭球面或球面上，则相应的 ω_3 光均能被合成并出射晶体，并且由于它们的频率和波矢大小相等而相位差恒定（即使 ω_3 光都为 e 光或 o 光，偏振和极化方向也因波矢方向的不同而不一定相同），进而两两相干，因此生成的所有 ω_3 光再经振幅调制后均可作为还原的倍频物光波出射晶体，生成非线性全息图像。

当然，式(2.69)只是最简单的对二阶非线性系数的调制方法，也有更多更复杂的调制方法，只要能够人工地造成二阶非线性系数随空间的起伏的方法，均有潜力实现准相位匹配。接下来对三维非线性光子晶体的二阶非线性系数的设计和调制方法就属于其中之一。

2.3 利用三维非线性光子晶体实现傅里叶全息成像

正如绪论 1.2.1.3 节所言，利用三维非线性光子晶体实现非线性全息成像的过程中，由于非线性全息的特殊干涉机制即倍频极化光作为重现光和参考光均无法出射晶体，因此三维光子晶体必须既是非线性效应发生的场所，同时也是记录体积全息图的介质即体全息板；又因非线性全息属于计算全息，因此没有记录干涉图样过程，只有重现倍频物光波的过程，所以在实验之前，对三维非线性光子晶体的结构设计必须同时兼顾非线性光学效应与全息成像，即所设计的结构必须能先有较高效率实现基频光到倍频光的转换，同时不同波矢的倍频光的振幅会因不同波矢失配量所对应的不同倒格矢所对应的不同倒格波的不同振幅而有所空间分布，并还原出逆向设计出的二维图像，这样才能实现整个非线性全息成像过程。

2.3.1 三维光子晶体中的准相位匹配机制

由于非线性光学效应的强度直接决定了倍频极化光的强度，而倍频极化波源的强度配合相位匹配程度又决定了倍频光的强度，而倍频光的强度即决定了还原出的倍频物光的强度，进而决定了远场所成的二维图像的整体明暗程度，因此对三维光子晶体的结构设计必须首先同时拥有较高的非线性系数和相位匹配程度。

2.3.1.1 为提高转换效率所采取的晶体放置方位和入射光种类选择以及偏振方向

而三维铌酸锂晶体的非线性系数中 d_{33} 是最大的，但之前提到过，要想让 d_{33} 出现在 d_{eff} 中，必须让参与混频的所有光的极化方向大致相同，准确地说，晶体内 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 的偏振方向完全相同且与光轴完全重叠，而由于 o 光电矢量恒垂直于光轴，只有 e 光的电矢量在其主平面内而才有可能平行于光轴，则 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 必须均为 e 光，此时 $\tilde{\mathbf{E}}_1 = \tilde{\mathbf{E}}_{1z}$, $\tilde{\mathbf{E}}_2 = \tilde{\mathbf{E}}_{2z}$ ，二者的其他分量全为零，代入式(2.34)得：

$$\begin{bmatrix} P_x^{(2)}(\omega_1 + \omega_2) \\ P_y^{(2)}(\omega_1 + \omega_2) \\ P_z^{(2)}(\omega_1 + \omega_2) \end{bmatrix} = 4\epsilon_0 \begin{bmatrix} d_{13} \\ d_{23} \\ d_{33} \end{bmatrix} \cdot E_z(\omega_1)E_z(\omega_2) \quad (2.85)$$

可见此时可以最大程度地利用 d_{33} : 极化时不将能量分散开来以同时利用上其他二阶非线性系数分量，而将 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 的偏振方向完全集中于光轴即 z 向，以期将 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 的振幅和能量集中于 z 向，丢掉其它偏振分量和二阶非线性系数分量，z 向完全只通过 d_{33} 而非六个分量，进行转换并极化，由于入射光的偏振方向充分利用了 d_{33} 这一最大非线性系数，因此极化所得的 $\tilde{\mathbf{P}}_3^{(2)}$ 振幅和能量也最高。并且此时由于 d_{33} 最大， $\tilde{\mathbf{P}}_3^{(2)}$ 的 z 分量也最大，因此其极化方向也大致朝向光轴方向（这里的 z 因牵涉到非线性系数的角标而均代表光轴方向），而由于在傍轴近似下， $\tilde{\mathbf{E}}_3$ 波矢与 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 波矢近似同向，则其偏振方向由于也近似垂直于其波矢方向，且由于是 e 光^①而电矢量位于主平面内，

^① $\tilde{\mathbf{E}}_3$ 是 e 光的原因：基频光入射晶体后因折射率椭球+两种可能的偏振方向所导致的两个折射率曲面所导致的两个折射率、两个无非线性波源项的波动方程、两个定振幅 o 光 e 光解，因此当两束光入射晶体后，在非线性过程发生之前本质上由四条光，各自的 e 光和 o 光有两对。两对 o 光 e 光之间有 4 种混频组合，而同一对内每一对均有 3 种混频组合，一共 $4+3+3=10$ 种混频组合，每一种混频组合中两个混频成分，即所谓的选择公式中的 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 为 o 光或 e 光，此时的 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 已经不再是两束入射光了，而是两束入射光在晶体内分出的四种成分中可被重复选择的两者。——现只考虑其中某一束入射光在晶体内的 e 光，被重复选择为 $\tilde{\mathbf{E}}_1$ 和 $\tilde{\mathbf{E}}_2$ 的指代对象，即考虑晶体内 4 束光束中某一束光束的倍频过程，那么在非线性过程发生并生成了倍频极化光后，倍频极化光开始通过有非线性波源的波动方程生成随传播距离增加振幅

(下页续)

所以 $\tilde{\mathbf{E}}_3$ 与 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 也近似同向，则 $\tilde{\mathbf{E}}_3$ 也大致朝向光轴方向，因此此时 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2, \tilde{\mathbf{P}}_3^{(2)}, \tilde{\mathbf{E}}_3$ 四者与光轴大致同向（前两者是完全重合）。而由于对 e 光而言电矢量偏振方向与波矢只是大致垂直，则此时光轴也与 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 的波矢也大致垂直。以上便意味着无法使用双折射实现相位匹配，只能采用准相位匹配技术。

2.3.1.2 三维电磁各向同性非线性电介质中的定态波动方程的标量化

一般正空间的非线性全息采用倍频极化光作为重现光，而在三维非线性全息中直接对倍频光的振幅进行调制，因此非线性过程仍只考虑到二阶非线性效应中的二次谐波生成过程，当然不失一般性地，也一并考虑三维情形下的和频波生成过程；又由于同样不考虑双折射，则波动方程因铌酸锂晶体的电各向异性所引入的两种等效标量折射率的问题也暂时不予考虑，因此可沿用三维空间中电磁各向同性非线性电介质中的定态波动方程即式(2.33)。但现先将三维情形下仍成立的 $\nu_3^2 = 1/\left[\mu_3^{(1)}\epsilon_3^{(1)}\right] = \omega_3^2/k_3^2$ 所对应的 $k_3^2 = \omega_3^2\mu_3^{(1)}\epsilon_3^{(1)}$ 代入式(2.33)左侧，右侧保持原样，得：

$$\nabla^2 \mathbf{E}_3(\mathbf{r}) + k_3^2 \mathbf{E}_3(\mathbf{r}) = -\frac{k_3^2}{\epsilon_3^{(1)}} \mathbf{P}_3^{(2)}(\mathbf{r}) \quad (2.86)$$

然而即使在三维情形下，也不希望使用式(2.34)来描述完整却复杂的非线性过程，因此接下来仍会尝试着将式(2.33)中各量标量化的同时，又不降低波动方程的维度。

为此，同样忽略铌酸锂晶体的双折射、认为铌酸锂晶体的二阶非线性系数的色散可以忽略以至于克莱曼对称性对二阶非线性系数仍不失效、但保留铌酸锂晶体的一阶非线性系数的正常色散（毕竟这是补偿波矢失配的目标），并只考虑特定入射的几何场景和选定的坐标系。将所选坐标系下的三维铌酸锂晶体的二阶非线性系数的各分量代入式(2.85)中，并两边取大小（不是取模，要保留场的相位部分），得到 SFG 和 SHG 的标量极化光的复振幅表达式：

的倍频光，那是不是两个 e 光通过有源波动方程生成的倍频光就是 e 光呢？不是，因为若这样的话，则两个 o 光生成的也是倍频 o 光，则表 2.3 中的 ooe 类型的双折射相位匹配就实现了；何况只要是通过各向异性晶体中波动方程生成的波，不论是有源还是无源的，由于左侧均含有一阶非线性系数，而必定有两种折射率，因此任何波动方程都有两个，以至于有两个解，因此 e 光通过波动方程生成的倍频光有两种，倍频 e 光和倍频 o 光，但由于倍频 o 光为与基频光 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 能流方向相同，则倍频 o 光的波矢球面上的波矢必须与 $\mathbf{k}_1 + \mathbf{k}_2$ 不同，因此若考虑波矢与 $\mathbf{k}_1 + \mathbf{k}_2$ 近似同向的 e 光和 o 光，则 o 光会因能流方向与 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 的不同而最终振幅不够大，而倍频 e 光能流方向与同为 e 光的基频光大致相同，则傍轴方向 e 光占大部分。

$$P_3^{(2)}(\mathbf{r}) = \varepsilon_0 \chi_{\text{eff}}^{(2)}(\mathbf{r}) \cdot E_1(\mathbf{r}) E_2(\mathbf{r}) = \varepsilon_0 \chi_{\text{eff}}^{(2)}(\mathbf{r}) \cdot A_1(\mathbf{r}) A_2(\mathbf{r}) e^{i(k_1+k_2) \cdot \mathbf{r}} =: p_3^{(2)}(\mathbf{r}) e^{i(k_1+k_2) \cdot \mathbf{r}} \quad (2.87)$$

其中 $P^{(2)}(\omega_3) = |\mathbf{P}_3^{(2)}| \arg[\mathbf{P}_3^{(2)}]$ 、 $E_1(\mathbf{r}) = |\mathbf{E}_1(\mathbf{r})| \arg[\mathbf{E}_1(\mathbf{r})]$ 、 $E_2(\mathbf{r}) = |\mathbf{E}_2(\mathbf{r})| \arg[\mathbf{E}_2(\mathbf{r})]$ 且

$$\chi_{\text{eff}}^{(2)}(\mathbf{r}) = \begin{cases} 2d_{\text{eff}}(x, y, z), & \omega_1 = \omega_2 \\ 4d_{\text{eff}}(x, y, z), & \omega_1 \neq \omega_2 \end{cases} \quad (2.88)$$

是经过三维设计并调制后的含空二阶非线性系数。

同样对式(2.86)剩余物理量即 $\mathbf{E}_3(\mathbf{r})$ 保留相位部分地标量化 $E_3(\mathbf{r}) = |\mathbf{E}_3(\mathbf{r})| \arg[\mathbf{E}_3(\mathbf{r})] = A_3(\mathbf{r}) e^{ik_3 \cdot \mathbf{r}}$ ，并将其与非线性波源项即式(2.87)代入式(2.86)，得到三维标量波动方程：

$$\nabla^2 E_3(\mathbf{r}) + k_3^2 E_3(\mathbf{r}) = -\frac{k_3^2}{\varepsilon_3^{(1)}} p_3^{(2)}(\mathbf{r}) e^{i(k_1+k_2) \cdot \mathbf{r}} \quad (2.89)$$

2.3.1.3 倍轴近似和缓变振幅近似条件下的标量波动方程及三个耦合波方程

当由正常色散引起的波矢失配量及补上该失配量的倒格矢相比于倍频极化波的波矢不大时，或者说在倍轴近似条件下，参与混频的三波以及非线性极化波，四个波的波矢大致同向，则假设 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2, \tilde{\mathbf{P}}_3^{(2)}, \tilde{\mathbf{E}}_3$ 四者的波矢方向均大致沿 z 轴正向，即虽不一定与 z 轴严格重合（为增大 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 而需利用上 d_{33} ，则光轴在这里近似垂直于 z 轴），但满足倍轴近似条件即 $k_3 \approx k_{3z}$ 、 $k_{3x}, k_{3y} \approx 0$ （最好写作 $k_{3x}, k_{3y} \ll k_{3z}$ 这种相对关系，之后有了结构倒格矢后可写作更准确的关系即 $\tan \delta = |\mathbf{G}_{m_x, m_y, m_z} \times (\mathbf{k}_1 + \mathbf{k}_2)| / |\mathbf{k}_1 + \mathbf{k}_2|^2 \ll 1$ ），这么做的原因并不是要用它来化简接下来的波动方程，反而接下来对波动方程的操作不会用到倍轴近似条件，而是因为全都是 e 光的 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2, \tilde{\mathbf{E}}_3$ 三者的波矢方向最好近似同向，否则若 $k_{3x}, k_{3y} \gg k_{3z}$ 则 \mathbf{k}_3 与 $\mathbf{k}_1 + \mathbf{k}_2$ 偏离程度较大，则 $\tilde{\mathbf{S}}_3$ 与 $\tilde{\mathbf{S}}_1 + \tilde{\mathbf{S}}_2$ 偏离较大，以至于向前传播到一定距离后场 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 与场 $\tilde{\mathbf{E}}_3$ 的交叠程度减弱，导致 $\tilde{\mathbf{E}}_3$ 振幅提前停止增长。在倍轴近似的前提下，方程(2.89)左侧的 laplace 算符和波矢可分解为传播方向和上传播方向的两部分：

$$\nabla_T^2 E_3(\mathbf{r}) + \frac{\partial^2 E_3(\mathbf{r})}{\partial z^2} + k_{3z}^2 E_3(\mathbf{r}) + (k_3^2 - k_{3z}^2) E_3(\mathbf{r}) = -\frac{k_3^2}{\varepsilon_3^{(1)}} p_3^{(2)}(\mathbf{r}) e^{i(k_1+k_2) \cdot \mathbf{r}} \quad (2.90)$$

其中，左侧的 z 向部分可通过类似式(2.40)的过程得：

$$\frac{\partial^2 E_3(\mathbf{r})}{\partial z^2} + k_{3z}^2 E_3(\mathbf{r}) = \frac{\partial^2 A_3(\mathbf{r})}{\partial z^2} e^{ik_3 \cdot \mathbf{r}} + 2 \frac{\partial A_3(\mathbf{r})}{\partial z} i k_{3z} e^{ik_3 \cdot \mathbf{r}} \quad (2.91)$$

同样设有三维情形下的 z 向的缓变振幅近似：

$$\left| \frac{\partial^2 A_3(\mathbf{r})}{\partial z^2} \right| \ll \left| k_{3z} \frac{\partial A_3(\mathbf{r})}{\partial z} \right| \quad (2.92)$$

于是将式(2.91-2.92)代入式(2.90)并两侧同除以 $e^{ik_3 \cdot r}$ 得：

$$\nabla_T^2 [A_3(\mathbf{r})e^{ik_3 \cdot r}] / e^{ik_3 \cdot r} + 2ik_{3z} \frac{\partial A_3(\mathbf{r})}{\partial z} + (k_3^2 - k_{3z}^2) A_3(\mathbf{r}) = -\frac{k_3^2}{\epsilon_3^{(1)}} p_3^{(2)}(\mathbf{r}) e^{i(k_1+k_2-k_3) \cdot r} \quad (2.93)$$

其中左侧 $\perp z$ 向的部分类似(2.91)地有

$$\nabla_T^2 [A_3(\mathbf{r})e^{ik_3 \cdot r}] / e^{ik_3 \cdot r} + (k_3^2 - k_{3z}^2) A_3(\mathbf{r}) = \frac{\partial^2 A_3(\mathbf{r})}{\partial x^2} + \frac{\partial^2 A_3(\mathbf{r})}{\partial y^2} + 2ik_{3x} \frac{\partial A_3(\mathbf{r})}{\partial x} + 2ik_{3y} \frac{\partial A_3(\mathbf{r})}{\partial y} \quad (2.94)$$

同样假设 x, y 向的缓变振幅条件成立，但暂时不作傍轴近似即 $k_{3x}, k_{3y} \approx 0$ ，则下面二阶导项可以扔掉，一阶导项得到保留：

$$\left| \frac{\partial^2 A_3(\mathbf{r})}{\partial x^2} \right| \ll \left| k_{3x} \frac{\partial A_3(\mathbf{r})}{\partial z} \right|, \quad \left| \frac{\partial^2 A_3(\mathbf{r})}{\partial y^2} \right| \ll \left| k_{3y} \frac{\partial A_3(\mathbf{r})}{\partial z} \right| \quad (2.95)$$

于是将式(2.94-2.95)代入式(2.93)，得：

$$2ik_{3x} \frac{\partial A_3(\mathbf{r})}{\partial x} + 2ik_{3y} \frac{\partial A_3(\mathbf{r})}{\partial y} + 2ik_{3z} \frac{\partial A_3(\mathbf{r})}{\partial z} = 2ik_3 \cdot \nabla A_3(\mathbf{r}) = -\frac{k_3^2}{\epsilon_3^{(1)}} p_3^{(2)}(\mathbf{r}) e^{i(k_1+k_2-k_3) \cdot r} \quad (2.96)$$

便有了第一个耦合波方程：

$$\begin{aligned} \frac{dA_3(\mathbf{r})}{\hat{\mathbf{k}}_3 \cdot d\mathbf{l}} &= \hat{\mathbf{k}}_3 \cdot \nabla A_3(\mathbf{r}) = \frac{ik_3}{2\epsilon_3^{(1)}} p_3^{(2)}(\mathbf{r}) e^{i\Delta\mathbf{k} \cdot \mathbf{r}} \xrightarrow[k_3=\frac{|\omega_3|\sqrt{\mu_{r3}^{(1)}} n_{3e}}{\epsilon_3^{(1)}=\epsilon_0 n_{3e}^2 c}]{} \frac{i|\omega_3| \sqrt{\mu_{r3}^{(1)}}}{2\epsilon_0 n_{3e} c} p_3^{(2)}(\mathbf{r}) e^{i\Delta\mathbf{k} \cdot \mathbf{r}} \\ &\xrightarrow[p_3^{(2)}(\mathbf{r})=2\epsilon_0 \chi_{eff}^{(2)}(\mathbf{r}) \cdot A_1(\mathbf{r}) A_2(\mathbf{r})]{} \frac{i|\omega_3| \sqrt{\mu_{r3}^{(1)}} \chi_{eff}^{(2)}(\mathbf{r})}{2n_{3e} c} A_1(\mathbf{r}) A_2(\mathbf{r}) e^{i\Delta\mathbf{k} \cdot \mathbf{r}} \end{aligned} \quad (2.97)$$

其中 $\hat{\mathbf{k}}_3 \cdot \nabla A_3(\mathbf{r})$ 为标量场 $A_3(\mathbf{r})$ 沿 $\hat{\mathbf{k}}_3$ 方向的方向导数 $dA_3(\mathbf{r}) / (\hat{\mathbf{k}}_3 \cdot d\mathbf{l})$ ，其中 $d\mathbf{l}$ 为朝空间任意方向的微小矢量段，不需要与 $\hat{\mathbf{k}}_3$ 平行； $\hat{\mathbf{k}}_3 \cdot d\mathbf{l} = d\mathbf{l}_{\hat{\mathbf{k}}_3}$ 便是沿 $\hat{\mathbf{k}}_3$ 方向的微小标量段； $dA_3(\mathbf{r}) = (\hat{\mathbf{k}}_3 \cdot d\mathbf{l}) \hat{\mathbf{k}}_3 \cdot \nabla A_3(\mathbf{r}) = d\mathbf{l}_{\hat{\mathbf{k}}_3} \hat{\mathbf{k}}_3 \cdot \nabla A_3(\mathbf{r}) = d\mathbf{l}_{\hat{\mathbf{k}}_3} \cdot \nabla A_3(\mathbf{r})$ 即标量场 $A_3(\mathbf{r})$ 与朝 $\hat{\mathbf{k}}_3$ 方向的微小矢量位移 $d\mathbf{l}_{\hat{\mathbf{k}}_3}$ 相对应的标量微分量。由于 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2, \tilde{\mathbf{P}}_3^{(2)}, \tilde{\mathbf{E}}_3$ 四者的波矢方向均大致沿 z 轴正向，则对参与混频的两个入射波也有类似的傍轴近似条件 $k_1 \approx k_{1z}, k_{1x}, k_{1y} \approx 0$ 、 $k_2 \approx k_{2z}, k_{2x}, k_{2y} \approx 0$ 成立（但推导过程中仍不会使用它），而又因克莱曼对称性成立使 $\chi_{eff}^{(2)}(\mathbf{r})$ 不含频率，则同理可导出其余两个标量形式的耦合波方程：

$$\frac{dA_1(\mathbf{r})}{\hat{\mathbf{k}}_1 \cdot d\mathbf{l}} = \hat{\mathbf{k}}_1 \cdot \nabla A_1(\mathbf{r}) = \frac{i|\omega_1| \sqrt{\mu_{r1}^{(1)}} \chi_{eff}^{(2)}(\mathbf{r})}{2n_{1e} c} A_3(\mathbf{r}) A_2^*(\mathbf{r}) e^{-i\Delta\mathbf{k} \cdot \mathbf{r}} \quad (2.98)$$

$$\frac{dA_2(\mathbf{r})}{\hat{\mathbf{k}}_2 \cdot d\mathbf{l}} = \hat{\mathbf{k}}_2 \cdot \nabla A_2(\mathbf{r}) = \frac{i|\omega_2| \sqrt{\mu_{r2}^{(1)}} \chi_{eff}^{(2)}(\mathbf{r})}{2n_{2e} c} A_3(\mathbf{r}) A_1^*(\mathbf{r}) e^{-i\Delta\mathbf{k} \cdot \mathbf{r}} \quad (2.99)$$

2.3.1.4 二阶非线性系数经三维周期性空间调制下的准相位匹配机制

其中，设经空间调制的 $\chi_{eff}^{(2)}(\mathbf{r})$ 及其调制函数 $B(\mathbf{r}) = B(\Lambda_x, \Lambda_y, \Lambda_z; \mathbf{r})$ 在 x, y, z 三个方向上均是周期函数，即使没有周期，也可以晶体内部调制的空间范围在 x, y, z 三个方向上的长度为周期，用 B-K 边界条件将其延拓成周期函数，并设结构的正空间周期为 $\Lambda_x, \Lambda_y, \Lambda_z$ ，则二阶非线性系数及其调制函数可展为三维复数形式傅里叶级数：

$$\chi_{eff}^{(2)}(\mathbf{r}) = \chi_{eff}^{(2)} \cdot B_{\Lambda_x, \Lambda_y, \Lambda_z}(\mathbf{r}) = \chi_{eff}^{(2)} \cdot \sum_{m_x, m_y, m_z=-\infty}^{+\infty} C_{m_x, m_y, m_z} e^{iG_{m_x, m_y, m_z} \cdot \mathbf{r}} \quad (2.100)$$

其中倒格波的振幅

$$\begin{aligned} C_{m_x, m_y, m_z} &= \mathcal{F}\left\{B_{\Lambda_x, \Lambda_y, \Lambda_z}(\mathbf{r})\right\} = \frac{1}{\Lambda_x \Lambda_y \Lambda_z} \int_{-\Lambda_x/2}^{\Lambda_x/2} \int_{-\Lambda_y/2}^{\Lambda_y/2} \int_{-\Lambda_z/2}^{\Lambda_z/2} B_{\Lambda_x, \Lambda_y, \Lambda_z}(\mathbf{r}) e^{-i(G_{m_x}x + G_{m_y}y + G_{m_z}z)} dx dy dz \\ &= \frac{1}{\Omega} \iiint_{\Omega} B_{\Omega}(\mathbf{r}) e^{-iG_{m_x, m_y, m_z} \cdot \mathbf{r}} d^3 r \end{aligned} \quad (2.101)$$

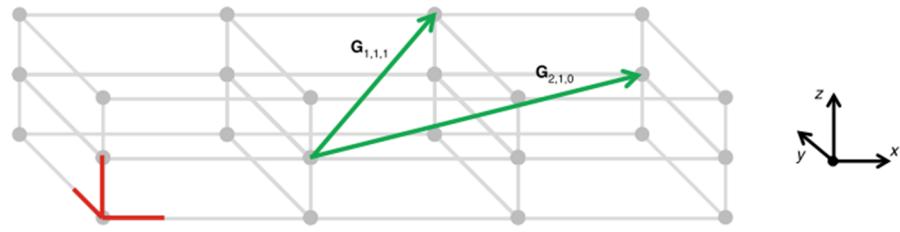


图 2.9 三维准相位匹配过程中弥补各个方向的波矢失配的不同级次倒格矢^[13]

其中， $\Omega = \Lambda_x \Lambda_y \Lambda_z$ 为三维光子晶体正空间中的原胞/晶胞即周期性结构的体积大小， $d^3 r = dx dy dz$ ，且

$$\mathbf{G}_{m_x, m_y, m_z} = \mathbf{G}_{m_x} + \mathbf{G}_{m_y} + \mathbf{G}_{m_z} = \frac{2\pi m_x}{\Lambda_x} \mathbf{i} + \frac{2\pi m_y}{\Lambda_y} \mathbf{j} + \frac{2\pi m_z}{\Lambda_z} \mathbf{k} \quad (2.102)$$

在之前，一维的调制函数是已知的且可写出表达式，其傅里叶变换所得的各单色倒格波的振幅在某些级有最大值，因此当时单纯为了提高出射的唯一一束和频光振幅，会选择依照该级倒格波所对应的倒格矢，来弥补波矢失配，并参照之设计一维的调制周期。但现在既要考虑非线性过程的出射倍频光振幅整体不能太低，又要考虑非线性全息成像即出射的多束倍频光的振幅起伏、分布问题，所以调制函数 $B_{\Omega}(\mathbf{r})$ 的设计并不需要也不能使其中一项倒格波的振幅 C_{m_x, m_y, m_z} 占主导地位，而是通过设计正空间的 $B_{\Omega}(\mathbf{r})$ 分布，以得到倒空间中各倒格波的振幅 C_{m_x, m_y, m_z} 分布，以及倒格子原胞尺寸和点阵分布即倒格矢分布，并同时实现多个波矢匹配过程并出射多束振幅不同的倍频光，以最终在距离晶体厘米级的光屏上实现非线性全息成像：

$$\chi_{eff}^{(2)}(\mathbf{r}) \approx \chi_{eff}^{(2)} \cdot C_{m_x, m_y, m_z} e^{i\mathbf{G}_{m_x, m_y, m_z} \cdot \mathbf{r}} := \chi_Q^{(2)} e^{i\mathbf{G}_{m_x, m_y, m_z} \cdot \mathbf{r}} \quad (2.103)$$

将其代入三个耦合波方程得：

$$\frac{dA_3(\mathbf{r})}{dl_{\hat{\mathbf{k}}_3}} = \frac{dA_3(\mathbf{r})}{\hat{\mathbf{k}}_3 \cdot dl} = \hat{\mathbf{k}}_3 \cdot \nabla A_3(\mathbf{r}) = \frac{i|\omega_3| \sqrt{\mu_{r3}^{(1)}} \chi_Q^{(2)}}{2n_{3e}c} A_1(\mathbf{r}) A_2(\mathbf{r}) e^{i\Delta k_Q \cdot \mathbf{r}} \quad (2.104)$$

$$\frac{dA_1(\mathbf{r})}{dl_{\hat{\mathbf{k}}_1}} = \frac{dA_1(\mathbf{r})}{\hat{\mathbf{k}}_1 \cdot dl} = \hat{\mathbf{k}}_1 \cdot \nabla A_1(\mathbf{r}) = \frac{i|\omega_1| \sqrt{\mu_{r1}^{(1)}} \chi_Q^{(2)}}{2n_{1e}c} A_3(\mathbf{r}) A_2^*(\mathbf{r}) e^{-i(\Delta k_Q - 2\mathbf{G}_{m_x, m_y, m_z}) \cdot \mathbf{r}} \quad (2.105)$$

$$\frac{dA_2(\mathbf{r})}{dl_{\hat{\mathbf{k}}_2}} = \frac{dA_2(\mathbf{r})}{\hat{\mathbf{k}}_2 \cdot dl} = \hat{\mathbf{k}}_2 \cdot \nabla A_2(\mathbf{r}) = \frac{i|\omega_2| \sqrt{\mu_{r2}^{(1)}} \chi_Q^{(2)}}{2n_{2e}c} A_3(\mathbf{r}) A_1^*(\mathbf{r}) e^{-i(\Delta k_Q - 2\mathbf{G}_{m_x, m_y, m_z}) \cdot \mathbf{r}} \quad (2.106)$$

为实现准相位匹配，需有 $\Delta k_Q = \Delta k + \mathbf{G}_{m_x, m_y, m_z} = \mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 + \mathbf{G}_{m_x, m_y, m_z} = \mathbf{0}$ ，即需满足

$$\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{G}_{m_x, m_y, m_z} = \mathbf{k}_3 \quad (2.107)$$

此即“补上”波矢失配的含义：由于介质的正常色散，导致目标波矢 \mathbf{k}_3 比已有波矢之和 $\mathbf{k}_1 + \mathbf{k}_2$ 大，所以要通过周期性结构的倒格波的倒格矢 $\mathbf{G}_{m_x, m_y, m_z}$ 补上这一段尚欠缺的失配的波矢。将其写为分量形式即有，所设计的调制后的三维分布的二阶非线性系数的级数展开式中振幅占优的倒格波的倒格矢 $\mathbf{G}_{m_x, m_y, m_z}$ 必须同时满足下述 x, y, z 三个方向的波矢匹配条件：

$$\begin{cases} k_{1x} + k_{2x} + G_{m_x, m_y, m_z; x} = k_{3x} \\ k_{1y} + k_{2y} + G_{m_x, m_y, m_z; y} = k_{3y} \\ k_{1z} + k_{2z} + G_{m_x, m_y, m_z; z} = k_{3z} \end{cases} \quad (2.108)$$

2.3.2 结构倒格波的波矢与振幅和远程成像的关系

2.3.2.1 所有终点落在 \mathbf{k}_3 的旋转椭球波矢曲面上的倒格矢的级次下的和频光的波矢失配量均为零

但由于波动方程化为定态之初，只限制了和频光频率为 ω_3 ，没有限制其波矢 \mathbf{k}_3 的具体细节，只因介质的折射率确定了而导致该频率 ω_3 下，和频光 $\tilde{\mathbf{E}}_3$ 的波矢 \mathbf{k}_3 的大小才确定了： $k_3 = |\omega_3|/v_3 = |\omega_3|n_3/v_3 = |\omega_3|\sqrt{\mu_{r3}^{(1)}}n_{3e}/c$ ，又因和频光 $\tilde{\mathbf{E}}_3$ 为 e 光，因此和频光 $\tilde{\mathbf{E}}_3$ 的波矢 \mathbf{k}_3 满足：

$$|\mathbf{k}_3| = \frac{|\omega_3|\sqrt{\mu_{r3}^{(1)}}n_{3ee}}{c} \quad (2.109)$$

也就是说只限制了和频光 $\tilde{\mathbf{E}}_3$ 的波矢 \mathbf{k}_3 的大小，即起点为 e 光波矢曲面中心的 \mathbf{k}_3 ，其终点必须落在 e 光的折射率曲面(C107-C108)即

$$\frac{1}{n_{3ee}^2} = \frac{\sin^2 \theta_{3e}}{n_{3e\parallel}^2} + \frac{\cos^2 \theta_{3e}}{n_{3e\perp}^2}, \quad n_{3ee}(\theta_{3e}) = \frac{n_{3e\parallel} n_{3e\perp}}{\sqrt{n_{3e\perp}^2 \sin^2 \theta_{3e} + n_{3e\parallel}^2 \cos^2 \theta_{3e}}} \quad (2.110)$$

所确定的旋转椭球波矢曲面上（对于傍轴近似以及光轴近似垂直于波矢的情况下，波矢集中落在负单轴晶体的 e 光波矢曲面的短轴方向上的小许球面度所对应的有些许曲率的圆薄片上）：

$$|\mathbf{k}_3| = \frac{|\omega_3| \sqrt{\mu_{r3}^{(1)}}}{c} \cdot \frac{n_{3e\parallel} n_{3e\perp}}{\sqrt{n_{3e\perp}^2 \sin^2 \theta_{3e} + n_{3e\parallel}^2 \cos^2 \theta_{3e}}} \quad (2.111)$$

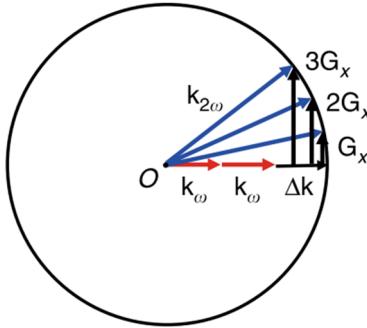


图 2.10 对倍频波与和频波的波矢只有大小限制而无方向限制^[14]

除此之外，对波矢 \mathbf{k}_3 的方向均没有限制。这意味着当极化光波矢 $\mathbf{k}_1 + \mathbf{k}_2$ 与和频光波矢 \mathbf{k}_3 的起点均置于旋转椭球波矢面(2.111)的对称中心时，对于任意级次的，起点在极化光波矢量 $\mathbf{k}_1 + \mathbf{k}_2$ 终点的倒格矢 $\mathbf{G}_{m_x, m_y, m_z}$ ，若其终点恰好落在(2.111)所确定的 k_3 曲面上，则其所对应的拥有波矢 $\mathbf{k}_{3; m_x, m_y, m_z} = \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{G}_{m_x, m_y, m_z}$ 的和频光 $\tilde{\mathbf{E}}_{3; m_x, m_y, m_z}$ 均因满足波矢匹配而可以随传播距离而振幅持续增大。那么所有满足式(2.111)的 $\mathbf{G}_{m_x, m_y, m_z}$ 都将被利用上，并同时出射与这些倒格矢相匹配的 $\mathbf{k}_{3; m_x, m_y, m_z} = \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{G}_{m_x, m_y, m_z}$ 所对应的和频光 $\tilde{\mathbf{E}}_{3; m_x, m_y, m_z}$ 们。

2.3.2.2 所有满足波矢匹配条件而出射的不同级次的和频光振幅均正比于相应级次的倒格波振幅

而之后会给出这些和频光因各自的振幅正比于相应级次的倒格矢 $\mathbf{G}_{m_x, m_y, m_z}$ 所对应的倒格波 $C_{m_x, m_y, m_z} e^{i\mathbf{G}_{m_x, m_y, m_z} \cdot \mathbf{r}}$ 的振幅 C_{m_x, m_y, m_z} （目前暂可从式(2.79)可见一斑），因此当各级次的倒格波的振幅系数 C_{m_x, m_y, m_z} 们因 $\chi_{eff}^{(2)}(\mathbf{r})$ 的具体调制情况有相应的三维的谱分布的同时，各符合波矢匹配而振幅可持续增长并可出射晶体的和频波 $\tilde{\mathbf{E}}_{3; m_x, m_y, m_z}$ 的振幅 $A_{3; m_x, m_y, m_z}(\mathbf{r})$ ，也将受到 $\chi_{eff}^{(2)}(\mathbf{r})$ 的调制函数的影响。

那么对 $\chi_{eff}^{(2)}(\mathbf{r})$ 的调制情况，将影响满足 $\mathbf{k}_{3; m_x, m_y, m_z} = \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{G}_{m_x, m_y, m_z}$ 的倒格波级次和具体的倒格矢 $\mathbf{G}_{m_x, m_y, m_z}$ 方向、各级和频波的波矢 $\mathbf{k}_{3; m_x, m_y, m_z}$ 方向、进而影响各级和频波的

能流 $\tilde{\mathbf{S}}_{3;m_x,m_y,m_z}$ 方向和出射方向，同时也会影响相应级次的和频波 $\tilde{\mathbf{E}}_{3;m_x,m_y,m_z}$ 的振幅 $A_{3;m_x,m_y,m_z}(\mathbf{r})$ 大小。

这样一来，在一定的球面度内，不仅团簇着出射的 $\tilde{\mathbf{E}}_{3;m_x,m_y,m_z}$ 的波矢和能流，在方向上有一定的分布， $\tilde{\mathbf{E}}_{3;m_x,m_y,m_z}$ 的振幅，也会随着不同波矢和能流方向，所对应的不同级次，而有一定的分布。同时，各级 $\tilde{\mathbf{E}}_{3;m_x,m_y,m_z}$ 因频率相同又是相干的，因此在对目标二维图案进行逆向设计完毕所得的 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 下，从晶体中出射的各级因满足波矢匹配而被筛选出来，并且振幅也归根结底被 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 调制的 $\{\tilde{\mathbf{E}}_{3;m_x,m_y,m_z}\}$ ，便可最终落在观察屏上并显示出设计好的二维全息图。这便是三维非线性光子晶体在非线性全息成像领域的应用之一。

下面将给出不同级次的和频光振幅表达式。仿照式(2.54)，沿 $\hat{\mathbf{k}}_3$ 方向进行积分：

$$\begin{aligned} A_3(\mathbf{r}; \mathbf{L}_{\hat{\mathbf{k}}_3}) &= A_3(\mathbf{r}; (\mathbf{r} - \mathbf{L}_{\hat{\mathbf{k}}_3}) \cdot \hat{\mathbf{k}}_3 = 0) = A_3(L_{\hat{\mathbf{k}}_3} = \mathbf{r} \cdot \hat{\mathbf{k}}_3) = A_3(L_{\hat{\mathbf{k}}_3} = \mathbf{r} \cdot \hat{\mathbf{k}}_3) - A_3(0 = \mathbf{r} \cdot \hat{\mathbf{k}}_3) \\ &= \frac{i|\omega_3| \sqrt{\mu_{r3}^{(1)}} \chi_{\text{Q}}^{(2)}}{2n_{3e}c} \int_0^{L_{\hat{\mathbf{k}}_3}} \left[A_1(\mathbf{r}) A_2(\mathbf{r}) e^{i\Delta k_Q \cdot \mathbf{r}} \right] \Big|_{\mathbf{r} \cdot \hat{\mathbf{k}}_3 = L_{\hat{\mathbf{k}}_3}} dL_{\hat{\mathbf{k}}_3} \end{aligned} \quad (2.112)$$

其中当参与混频的两个入射光场到和频光群的转换效率不太高时，和频光集 $\{\tilde{\mathbf{E}}_{3;m_x,m_y,m_z}\}$ 从 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 中吸取的能量并不多，不会造成 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 振幅的显著下降，此时两个参与混频的入射光场振幅近似恒定， $A_1(\mathbf{r}) = E_{01}$ 、 $A_2(\mathbf{r}) = E_{02}$ 不含空，则有：

$$A_3(\mathbf{r}; \mathbf{L}_{\hat{\mathbf{k}}_3}) = \frac{i|\omega_3| \sqrt{\mu_{r3}^{(1)}} \chi_{\text{Q}}^{(2)}}{2n_{3e}c} E_{01} E_{02} \int_0^{L_{\hat{\mathbf{k}}_3}} e^{i\Delta k_Q \cdot L_{\hat{\mathbf{k}}_3}} dL_{\hat{\mathbf{k}}_3} = \frac{i|\omega_3| \sqrt{\mu_{r3}^{(1)}} \chi_{\text{Q}}^{(2)}}{2n_{3e}c} E_{01} E_{02} \frac{e^{i\Delta k_Q \cdot L_{\hat{\mathbf{k}}_3}} - 1}{i\Delta k_Q \cdot \hat{\mathbf{k}}_3} \quad (2.113)$$

其中

$$\int_0^{L_{\hat{\mathbf{k}}_3}} e^{i\Delta k_Q \cdot L_{\hat{\mathbf{k}}_3}} dL_{\hat{\mathbf{k}}_3} = \int_0^{L_{\hat{\mathbf{k}}_3}} e^{i\Delta k_Q \cdot L_{\hat{\mathbf{k}}_3}} dL_{\hat{\mathbf{k}}_3} = \frac{e^{i\Delta k_Q \cdot L_{\hat{\mathbf{k}}_3}} - 1}{i\Delta k_Q \cdot \hat{\mathbf{k}}_3} = \frac{e^{i\Delta k_Q \cdot L_{\hat{\mathbf{k}}_3}} - 1}{i\Delta k_Q \cdot \hat{\mathbf{k}}_3} = \frac{e^{i(L \cdot \hat{\mathbf{k}}_3) \Delta k_Q \cdot \hat{\mathbf{k}}_3} - 1}{i\Delta k_Q \cdot \hat{\mathbf{k}}_3} \quad (2.114)$$

仿照式(2.55-2.59)的过程，其中由于此处 $\chi_{\text{eff}}^{(2)}$ 中的 $\chi_{\text{eff}}^{(2)}$ 比原 d_{eff} 在数学形式上小 4 倍，因此振幅即(2.113)比(2.54)和(2.79)在数学形式上小 4 倍，于是 I_3 在数学形式上应比(2.59)以及(2.78)小 16 倍，其他参数或不变或被替换，得倍频或和频过程结束后的：

$$I_3 = \frac{\sqrt{\mu_{r1}^{(1)} \mu_{r2}^{(1)} \mu_{r3}^{(1)}}}{n_{1e} n_{2e} n_{3e}} \frac{\omega_3^2}{8\varepsilon_0 c^3} \left(\chi_{\text{Q}}^{(2)} \right)^2 \cdot I_1 I_2 \cdot L_{\hat{\mathbf{k}}_3}^2 \operatorname{sinc}^2 \left(\frac{\Delta k_Q \cdot L_{\hat{\mathbf{k}}_3}}{2} \right)^{[30]} \quad (2.115)$$

同样对于和频过程结束即 $L_{\hat{\mathbf{k}}_3} \geq L_{\hat{\mathbf{k}}_3}$ 后，有类似式(2.79)的：

$$|E_{03}| = |A_3(\mathbf{r}; \mathbf{L}_{\hat{\mathbf{k}}_3})| = \frac{\sqrt{\mu_{r3}^{(1)}} |\omega_3|}{2n_{3e}c} \chi_{\text{Q}}^{(2)} \cdot |E_{01}| |E_{02}| \cdot L_{\hat{\mathbf{k}}_3} \cdot \left| \operatorname{sinc} \left(\frac{\Delta k_Q \cdot L_{\hat{\mathbf{k}}_3}}{2} \right) \right| \quad (2.116)$$

将该解或准相位匹配 $\Delta\mathbf{k}_Q = \mathbf{0}$ 时的式(2.116)代入式(2.92)、(2.95)均成立，说明之前所假设的 x, y, z 方向上的缓变振幅近似也均成立。

当准相位匹配条件 $\Delta\mathbf{k}_{Q;m_x,m_y,m_z} = \Delta\mathbf{k}_{m_x,m_y,m_z} + \mathbf{G}_{m_x,m_y,m_z} = \mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_{3;m_x,m_y,m_z} + \mathbf{G}_{m_x,m_y,m_z} = \mathbf{0}$ 得到满足，同时相应级次的波矢 $\mathbf{k}_{3;m_x,m_y,m_z}$ 和倒格矢 \mathbf{G}_{m_x,m_y,m_z} 的终点落在 \mathbf{k}_3 波矢曲面的相应旋转椭球面上时，非线性全息过程表述为：

$$\left\{ \begin{array}{l} \left| E_{3;m_x,m_y,m_z}(\mathbf{r}; \mathbf{r} \cdot \hat{\mathbf{k}}_{3;m_x,m_y,m_z} = l_{\hat{\mathbf{k}}_{3;m_x,m_y,m_z}} \geq L_{\hat{\mathbf{k}}_{3;m_x,m_y,m_z}}) \right| = \frac{\sqrt{\mu_{r3}^{(1)}} |\omega_3| \chi_{eff}^{(2)}}{2n_{3e}c} C_{m_x,m_y,m_z} |E_{01}| |E_{02}| L_{\hat{\mathbf{k}}_{3;m_x,m_y,m_z}} \\ \left| \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{G}_{m_x,m_y,m_z} \right| = \frac{|\omega_3| \sqrt{\mu_{r3}^{(1)}}}{c} \cdot \frac{n_{3e//} n_{3e\perp}}{\sqrt{n_{3e\perp}^2 \sin^2 \theta_{3e;m_x,m_y,m_z} + n_{3e//}^2 \cos^2 \theta_{3e;m_x,m_y,m_z}}} \end{array} \right. \quad (2.117)$$

可见在对二阶非线性系数 $\chi_{eff}^{(2)}(\mathbf{r})$ 进行了某一具体的三维空间调制（式(2.100)）即 $\chi_{eff}^{(2)}(\mathbf{r}) = \chi_{eff}^{(2)} \cdot B_{\Lambda_x, \Lambda_y, \Lambda_z}(\mathbf{r}) = \chi_{eff}^{(2)} \cdot \sum_{m_x, m_y, m_z=-\infty}^{+\infty} C_{m_x, m_y, m_z} e^{i\mathbf{G}_{m_x, m_y, m_z} \cdot \mathbf{r}}$ 后，二式给出了在这样的 $\chi_{eff}^{(2)}(\mathbf{r})$ 分布下，通过非线性过程、和频过程、准相位匹配即倒格矢弥补波矢失配的筛选过程，所留下的和频光集 $\{\tilde{\mathbf{E}}_{3;m_x,m_y,m_z}\}$ 的波矢集 $\{\mathbf{k}_{3;m_x,m_y,m_z}\}$ 在倒空间中的一定的发散程度和球面度分布；而一式则同时给出了不同级次、不同波矢、不同出射方向的和频光的振幅大小。

值得注意的是，对于式(2.117)中任何一个物理量，若其角标含 m_x, m_y, m_z ，说明该物理量或者说该物理量的自变量们是倒空间的，即描绘该物理量所选用的坐标，只能由衍射级次 (m_x, m_y, m_z) 和倒空间的三基矢 $2\pi/\Lambda_x \mathbf{i}, 2\pi/\Lambda_y \mathbf{j}, 2\pi/\Lambda_z \mathbf{k}$ 所决定的倒格矢

$\mathbf{G}_{m_x, m_y, m_z} = 2\pi m_x / \Lambda_x \mathbf{i} + 2\pi m_y / \Lambda_y \mathbf{j} + 2\pi m_z / \Lambda_z \mathbf{k}$ 来描述，那么其中的 $E_{3;m_x,m_y,m_z}$ 连同其波矢 $\hat{\mathbf{k}}_{3;m_x,m_y,m_z}$ 由于只能写出二者关于倒空间坐标的函数关系，则均只能由倒格矢描述，并且因此 $E_{3;m_x,m_y,m_z}$ 的振幅所携带的信息，是二阶非线性系数的倒空间信息，即其频谱信息。同样， $E_{3;m_x,m_y,m_z}$ 的实际传播距离、有效倍频距离、其波矢与光轴的夹角即

$l_{\hat{\mathbf{k}}_{3;m_x,m_y,m_z}}, \mathbf{L}_{\hat{\mathbf{k}}_{3;m_x,m_y,m_z}}, \theta_{3e;m_x,m_y,m_z}$ 三者，也均只能由倒空间的坐标量描绘。只不过由于正空间与倒空间三基矢的方向对应同向且均为 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ ，且基矢大小有数量关系，而可以通过坐标系变换得到坐标变换关系，并用于两种坐标系下的坐标过渡，比如式(2.117)中的 $\mathbf{k}_1, \mathbf{k}_2$ 、光轴方向和相应折射率椭球，四者可通过坐标变换从正空间过渡到倒空间。

合理设计 $\chi_{eff}^{(2)}(\mathbf{r})$ 可同时对不同传播方向上的倍频光、和频光的振幅进行调制，便在理论上允许了三维光子晶体可进行倍频波、和频波的三维非线性全息成像，并同时给出了非线性全息成像的过程和对于给定目标图像，相关物理量的设计和计算方法。

3 三维非线性结构设计与全息成像仿真

上一章描述了在其他条件已知时，对于给定的二阶非线性系数分布函数 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ ，可通过考察终点落在 ω_3 光的波矢曲面上的倒格矢 $((m_x, m_y, m_z), \mathbf{G}_{m_x, m_y, m_z})$ ，以及 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 或其调制函数的傅里叶变换即 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 的角谱 $((m_x, m_y, m_z), C_{m_x, m_y, m_z})$ ，确定出射的和频光集中，每一条级次的和频光的波矢、波矢方向和不同方向的振幅大小分布，即 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 的频谱 $(\mathbf{G}_{m_x, m_y, m_z}, C_{m_x, m_y, m_z})$ 经横坐标在倒空间中的三维平移 $(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{G}_{m_x, m_y, m_z}, C_{m_x, m_y, m_z})$ 后所得的平移后的频谱 $(\mathbf{k}_{3; m_x, m_y, m_z}, C_{m_x, m_y, m_z})$ ，或者更准确地说是横坐标于倒空间平移、纵坐标于频谱空间拉伸后的频谱 $(\mathbf{k}_{3; m_x, m_y, m_z}, C_{m_x, m_y, m_z} \cdot L_{\hat{\mathbf{k}}_3})$ ，进而正向确定出射光在近场光屏或 CCD 上的近场衍射光强分布图，即 $(\hat{\mathbf{k}}_{3; m_x, m_y, m_z}, |E_3(\hat{\mathbf{k}}_{3; m_x, m_y, m_z})|)$ 、 $(\mathbf{r}(\hat{\mathbf{k}}_{3; m_x, m_y, m_z}), |E_3(\mathbf{r})|)$ 、 $((x, y, z)|_{\hat{\mathbf{k}}_{3; m_x, m_y, m_z}}, I_{3; x, y, z})$ ，实现三维非线性近场菲涅尔计算体全息。

而本章内容则是在其他条件已知的情况下，对远场目标物 $((x, y, z)|_{\hat{\mathbf{k}}_{3; m_x, m_y, m_z}}, I_{3; x, y, z})$ ，反向数值计算和设计出二阶非线性系数 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 在正空间的三维调制函数，并依据所得结果，在仿真软件内编写程序，通过多值化编码方法将计算结果转换为三维光子晶体的正空间实际结构，然后仍通过数值仿真查看对于所设计的结构，出射倍频光的远场成像图案是否与目标物匹配，如果匹配且成像质量较高，说明对 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 的计算结果正确，并且相应的编码法对于将 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 转换为三维光子晶体内实际刻写的正空间结构是有效的，经验证各步骤的结果均符合预期、方案设计正确且成熟后，便可对实际的三维光子晶体进行飞秒激光直写加工，并以基频光照射，以查看远场倍频光场的分布情况，继而在完成三维非线性全息过程的理论验证、数值模拟验证之后，完成三维非线性傅里叶计算体全息过程的实验验证。

3.1 三维 GS 算法：约束结构角谱为目标图像逆向获得 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 分布

3.1.1 基于三维光子晶体的远场傅里叶计算体全息的逆向过程原理

上一章未提到，倍频光集 $\{E_{3; m_x, m_y, m_z}\}$ 所携带的信息是三维光子晶体结构的频谱信息，而倒空间的频谱信息只能通过二维傅里叶变换才能在实空间中呈现出来，因此不能直接查看晶体后端面的倍频光强分布，而是需要看 $\{E_{3; m_x, m_y, m_z}\}$ 经过一次夫琅禾费衍射后的远场分布图案，此时 $\{E_{3; m_x, m_y, m_z}\}$ 的振幅分布才能在正空间中以光强分布的形式体现出来。因此对此只能采取傅里叶全息技术而不能采取菲涅尔全息。

由于远场光强和振幅分布图案是目标图案，因此倍频光集 $\{E_{3;m_x,m_y,m_z}\}$ 所携带的三维光子晶体的 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 的角谱的振幅分布也需要是目标图案的振幅分布，于是 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 本身的三维分布函数可以说是目标图案的振幅分布的傅里叶变换，但由于目标图案是二维的， $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 与其角谱均是三维的分布，接下来就需要完成从目标图案到 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 的角谱，再用到三维 GS 算法到 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 的转换。

3.1.2 波长匹配问题

在仿真的程序中，假设基频光沿 y 轴正向传播，且正空间中对二阶非线性系数的调制周期为 $\Lambda_x = 4, \Lambda_y = 3, \Lambda_z = 4$ （单位为 μm ），该结构是个 $x-o-z$ 面为正方形，但厚度即 y 方向较薄的小长方体，并且在实空间中一共有 $16 \times 16 \times 16$ 个这样的长方体，每个长方体中心的点阵坐标储存为一个三维数组 $(x, y, z) = (n_x \cdot \Lambda_x, n_y \cdot \Lambda_y, n_z \cdot \Lambda_z)$ ，其中三坐标每一个的取值有 16 个： $n_x, n_y, n_z = [-8, 7] \in \mathbb{Z}$ ，这该周期结构相应的倒格子点阵为 $(2\pi m_x / \Lambda_x, 2\pi m_y / \Lambda_y, 2\pi m_z / \Lambda_z)$ ，且也有 $m_x, m_y, m_z = [-8, 7] \in \mathbb{Z}$ ，每一个倒格点即代表光子晶体的一个倒格子，任意两个倒格点之间的倒格矢，相当于一个倒格子到另一个倒格子的平移矢量。

在这个 $64 \times 48 \times 64$ 的大长方体调制区域外，二阶非线性系数 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 的值不再改变，但在数学上采用 B-K 周期性边界条件将该调制区域的 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 分布情况延拓到正空间中晶体内的其余部分，使得一共有 $16 \times 16 \times 16$ 个调制周期为 $\Lambda'_x = 64, \Lambda'_y = 48, \Lambda'_z = 64$ 的实际调制区域，相应格波波矢终点构成的点阵为 $(2\pi l_x / \Lambda'_x, 2\pi l_y / \Lambda'_y, 2\pi l_z / \Lambda'_z)$ ，其中 $l_x, l_y, l_z = [-8, 7] \in \mathbb{Z}$ ，之所以称其为格波波矢构成的倒格点阵，是因该倒格点阵全分布在同一个倒格子内，且每个 $2\pi/\Lambda_x \times 2\pi/\Lambda_y \times 2\pi/\Lambda_z$ 的长方体倒格子内都有这样一个点阵，类似固体物理中的格波波矢全都位于同一布里渊区内，比倒格子的分布更密集。

正如 2.2 节一维准相位匹配中，采取振幅最大的 ± 1 级倒格波的倒格矢进行波矢匹配，以确定 $\chi_{\text{eff}}^{(2)}(z)$ 的一维调制周期一样，这里也需要先假定最终出射晶体并在远场成像的倍频光集 $\{E_{3;m_x,m_y,m_z}\}$ 是某一级次 (m_x, m_y, m_z) 的倒格矢进行波矢匹配的结果。这里说某一级次 (m_x, m_y, m_z) ，却又暗示该级次下的倍频光不止一束，是因为在做了 B-K 周期性边界条件将以 $\Lambda_x = 4, \Lambda_y = 3, \Lambda_z = 4$ 为周期的尺寸为 $64 \times 48 \times 64$ 的调制区域延拓后，式(2.117)中的波矢匹配条件会相应地修正为：

$$|\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{G}_{m_x, m_y, m_z} + \mathbf{q}_{l_x, l_y, l_z}| = \frac{|\omega_3| \sqrt{\mu_{r3}^{(1)}}}{c} \cdot \frac{n_{3e//} n_{3e\perp}}{\sqrt{n_{3e\perp}^2 \sin^2 \theta_{3e; m_x, m_y, m_z} + n_{3e//}^2 \cos^2 \theta_{3e; m_x, m_y, m_z}}} \quad (3.1)$$

$E_{3;m_x,m_y,m_z}$ 连同其波矢 $\hat{\mathbf{k}}_{3;m_x,m_y,m_z}$ 也需要修正为 $E_{3;m_x,m_y,m_z;l_x,l_y,l_z}$ 、 $\hat{\mathbf{k}}_{3;m_x,m_y,m_z;l_x,l_y,l_z}$ ，且还有：

$$\chi_{eff}^{(2)}(\mathbf{r}) = \chi_{eff}^{(2)} \cdot B_{\Lambda_x, \Lambda_y, \Lambda_z}(\mathbf{r}) = \chi_{eff}^{(2)} \cdot \sum_{m_x, m_y, m_z=-8}^7 \sum_{l_x, l_y, l_z=-8}^7 C_{m_x, m_y, m_z; l_x, l_y, l_z} e^{i(G_{m_x, m_y, m_z} + q_{l_x, l_y, l_z}) \cdot \mathbf{r}} \quad (3.2)$$

其中

$$\mathbf{q}_{l_x, l_y, l_z} = q_{l_x} \mathbf{i} + q_{l_y} \mathbf{j} + q_{l_z} \mathbf{k} = \frac{2\pi l_x}{\Lambda'_x} \mathbf{i} + \frac{2\pi l_y}{\Lambda'_y} \mathbf{j} + \frac{2\pi l_z}{\Lambda'_z} \mathbf{k} \quad (3.3)$$

为三维光子晶体格波波矢，或称为细节倒格矢。现假设采取 $(m_x, m_y, m_z) = (2, 1, 0)$ 级次的倒格矢 $\mathbf{G}_{2,1,0}$ ，并采取 $(l_x, l_y, l_z) = (l_x, 6, l_z)$ 级细节倒格矢 $\mathbf{q}_{l_x, 6, l_z}$ 即 y 向第六层细节倒格矢 ($q_{l_x, 6, l_z; y} = q_{l_y=6} = 2\pi l_y / \Lambda'_y = -0.41888$)，则相应的傅里叶系数 $C_{m_x, m_y, m_z; l_x, l_y, l_z} = C_{2,1,0; l_x, 6, l_z}$ 。

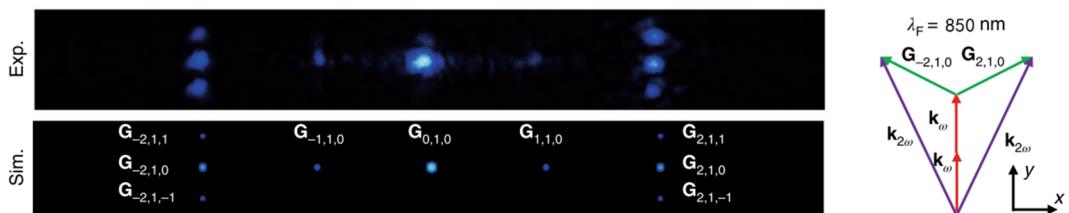


图 3.1 利用 2,1,0 级次倒格矢 $\mathbf{G}_{2,1,0}$ 补偿波矢失配示意图^[13]

接下来便是筛选匹配该衍射级的倒格矢和该层细节倒格矢的波长的过程，以使得该波长的基波所对应的倍频波的波矢 k_3 曲面将通过 $\mathbf{G}_{2,1,0} + \mathbf{q}_{l_x, 6, l_z}$ 附近：通过将范围位于 0.5~1.5 μm 之间（步长或增量为 0.01 μm）的波长一维数组，在温度 T=25°C 时，代入偏振方向平行于光轴的基频光入射晶体的物理场景下的铌酸锂的色散折射率函数 $n_e(\lambda, T)$ ，生成该范围内的波长所对应的基波波矢大小范围（也是一维数组），以及半波长、同样温度下，相应的倍频波波矢大小所构成的一维数组 $\{k_3\}$ 。

在细节波矢平面假定为 $\mathbf{q}_{l_x, 6, l_z}$ 的情况下，通过 matlab 的 find() 函数寻找这些不同波长的基频波和倍频波所对应的 y 方向的(细节)波矢失配量的最小值（即式 3.1 的 y 方向分量方程的右侧减左侧的绝对值的最小值），其所对应的基频波的波长，就是所采取的目标基频光的波长，在衍射级采取 $\mathbf{G}_{2,1,0}$ 和 $\mathbf{q}_{l_x, 6, l_z}$ 的情况下，数值仿真所给出的结果为 0.859 μm，可见三维光子晶体调制周期与匹配的波长处在同一个数量级，因此一方面说明以这样的周期调制的材料，称得上是三维光子晶体，另一方面也说明经这样调制的三维光子晶体理应具备操控相匹配的光波段的光场的能力。

在通过确定目标衍射级以确定了大致匹配的波长之后，才能继续采用该波长的基波所对应的倍频波的波矢 k_3 曲面进行接下来的计算。在 3.3 节，我们会更进一步地确定最匹配目标衍射级的波长为 0.850 μm，此处的 0.859 μm 只是大致匹配 $\mathbf{G}_{2,1,0}$ 倒格子内

的第六层细节倒格矢平面。

3.1.3 三维角谱分布函数的初始化

通过 3.1.2 节所确定的波长，只是在待选定的一维波长数组(0.501,0.502,...,1.500) μm 中，使得倍频光波矢 k_3 曲面最接近 y 向第 6 级细节倒格矢平面的基频光波长，其所对应的倍频光波矢 k_3 曲面并不完全落于第 6 级细节倒格矢平面即 $\mathbf{q}_{l_x,6,l_z}$ 上，原因有两个，一是其 y 向(细节)波矢失配量只是最小值，但该最小值不一定是零；二是倍频光波矢 k_3 曲面是个旋转椭球面，不可能与该平面重合。不过该曲面一定在相应细节倒格矢所在的倒格子，即倒格矢 $\mathbf{G}_{2,1,0}$ 所对应的倒格子内。

在 $\mathbf{q}_{l_x,6,l_z}$ 所在的 $\mathbf{G}_{2,1,0}$ 倒格子内，由于该倒格子在整个倒空间中所占体积很小，再加上铌酸锂晶体各波长下 n_{\perp}, n_{\parallel} 差别不大，导致 e 光的旋转椭球面没有那么扁或长，因此可将该 $\mathbf{G}_{2,1,0}$ 倒格子区域内的倍频 e 光的旋转椭球面近似为球面。

在这样的近似下，计算该 $\mathbf{G}_{2,1,0}$ 倒格子内不同 (q_{l_x}, q_{l_z}) 坐标下的倍频光波矢 \mathbf{k}_3 的 y 分量的大小 k_{3y} 的值，构成一个离散的二维分布函数 $k_{3y}(q_{l_x}, q_{l_z})$ 。接着用 matlab 的 sort 函数将每一个 (q_{l_x}, q_{l_z}) 下，一维数组 q_{l_y} 所对应的 $|k_{3y}(q_{l_x}, q_{l_z}) - q_{l_y}|$ 的值进行升序排列，并取其中绝对值最小的值所对应的 q_{l_y} ，即离曲面 $k_{3y}(q_{l_x}, q_{l_z})$ 最近的相应级次的细节倒格矢的 y 分量 q_{l_y} ，这样便找到了离 e 光的旋转椭球面最近的三维的离散细节倒格矢 $(q_{l_x}, q_{l_y}, q_{l_z})$ 的分布情况。

于是在 $\mathbf{G}_{2,1,0}$ 内的每一个 (q_{l_x}, q_{l_z}) 下，将这些细节倒格矢所对应的 $\chi_{eff}^{(2)}(\mathbf{r})$ 在倒空间的振幅 $C_{2,1,0;l_x,l_y,l_z}$ 置为目标 16×16 大小的图片（与 $l_x, l_z = [-8, 7] \in \mathbb{Z}$ 在 $x-o-z$ 方向上也为 16×16 的像素矩阵相对应）的相应 (q_{l_x}, q_{l_z}) 处的像素点的振幅。至此便完成了对三维角谱分布函数的初始值即初始三维分布情况的确定，注意 $C_{2,1,0;l_x,l_y,l_z}$ 是个由三维倒空间坐标描述的复数数组，只不过其初始化的值在 e 光旋转椭球面附近为目标图片的振幅，而在 $\mathbf{G}_{2,1,0}$ 内的其他 $(q_{l_x}, q_{l_y}, q_{l_z})$ 下的值，全为零，以至于该复数数组只是初始值全为实数罢，之后的 GS 算法的中间过程中它会被赋值为复数数组。

3.1.4 三维 GS 算法实现过程

3.1.4.1 二维 GS 算法的目的和操作流程

一般对一个实空间光场振幅分布，即一个实数二维数组的傅里叶变换，即其频谱或其傅里叶全息图的二维数组中每一个像素点的值，既有振幅部分又有相位部分，若只取其中的相位部分，振幅部分全置 1（全取其他常实值效果也类似），则再对该纯

相位全息图(POH)的进行傅里叶逆变换回到正空间物光场分布，并只取其振幅部分所还原得到的重建像，仍能得到目标物光场振幅分布的大致轮廓，但该过程中由于期间在将傅里叶全息图转换为纯相位傅里叶全息图时，“只取全息图的相位部分而直接扔掉了其振幅部分的信息”的操作，以及最终“只取纯相位全息图傅里叶逆变换回正空间的物光场的振幅而抛弃其相位”的操作，两次舍弃一半的信息，导致最终重建像质量不高，干涉散斑严重。

而典型的二维 GS 算法便针对目标实空间光场振幅分布，在经上述过程得到纯相位全息图的傅里叶逆变换所得的复数形式的物光场分布后，不直接取振幅得到重建像，而是将其振幅约束为目标图像振幅，相位不变，将所得的既有振幅又有相位的复数形式的物光场，作为新的输入再继续相同的第二轮操作，直到迭代次数升到一定程度（一般 50 次即可），使得最终重建像成像质量随着迭代次数的增加而上升的效果不再明显后，停止迭代并对最后一次迭代所得的复数形式的正空间物光场取振幅输出为重建像。

简言之，Gerchberg 和 Saxton 所提出的这个 GS 算法，目的是为了对于给定的正空间物光场，在迭次傅里叶变换(IFTA)的过程中，每次做了傅里叶变换后只取傅里叶全息图的相位，每次做了傅里叶逆变换后限制正空间光场的振幅，最终收敛到一个纯相位傅里叶全息图，其所还原得到的重建像的质量，虽比原像质量差一点，但比不迭代并只取相位的傅里叶全息图的傅里叶逆变换所得的重建像的质量要好。

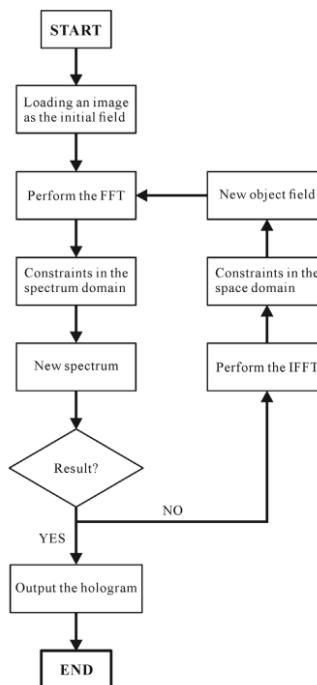


图 3.2 GS 算法即迭次傅里叶变换算法的流程图

3.1.4.2 三维 GS 算法的流程

三维 GS 算法与之类似，但目标振幅分布与其傅里叶全息图均是三维数组，并且目标振幅的三维分布在倒空间而非正空间中，且将其傅里叶逆变换作为振幅置 1 的纯相位三维数组，与通常意义上的 GS 算法中的傅里叶变换关系对的振幅约束对象恰好相反。在本仿真中，每次傅里叶变换后的结果为 $\mathbf{G}_{2,1,0}$ 倒格子内离散细节倒格矢 $(q_{l_x}, q_{l_y}, q_{l_z})$ 分布所对应的复角谱系数 $C_{2,1,0;l_x, l_y, l_z}$ ，它是一个关于倒空间三个方向的坐标的三维函数，其自变量是个 $16 \times 16 \times 16$ 的三维数组，接着将复角谱分布的振幅部分置为 3.1.3 节末的初始化 $C_{2,1,0;l_x, l_y, l_z}$ 实振幅三维近球面分布情况，并进行傅里叶逆变换回到 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 的实空间复函数分布，接着将 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 的振幅约束为 1 相位保持不变地转换为正空间的纯相位分布，于是按照这样的循环迭代 10000 次，取最后一次经由 $C_{2,1,0;l_x, l_y, l_z}$ 傅里叶逆变换所得的 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 复数分布的相位部分和其中的相位角，便得到正空间中 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 的复数纯相位分布和实数相位角分布，之后将用此实的相位角的三维分布对三维光子晶体进行多值化编码。

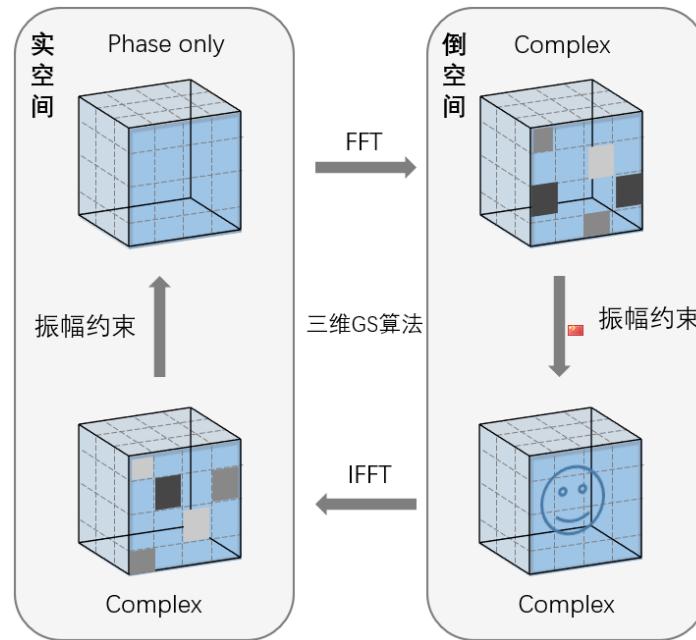


图 3.3 利用三维 GS 算法计算二阶非线性系数的三维相位分布

其中可以取最后一次迭代所得的 $C_{2,1,0;l_x, l_y, l_z}$ 复数数组的振幅，输出为 y 向排列的共 16 个不同 y 向细节倒格矢分量 q_{l_y} 所对应的 16 个 (q_{l_x}, q_{l_z}) 平面上的倒空间二维角谱分布图 $C_{2,1,0;l_x, l_y, l_z}$ ，查看三维 GS 算法的效果是否达到预期，即所设计的正空间的三维纯相位 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 分布的三维角谱在某一两个 (q_{l_x}, q_{l_z}) 平面上的二维分布，是否是目标图像的光场振幅分布。

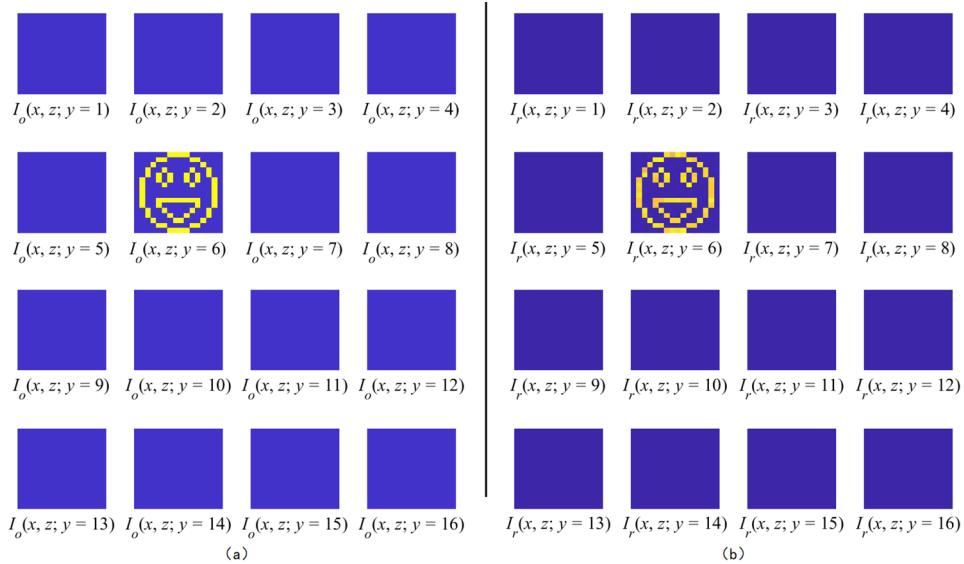


图 3.4 目标 $C_{2,1,0;l_x,l_y,l_z}$ 分布 (a) 和 GS 算法所得结构在倒空间的实际 $C_{2,1,0;l_x,l_y,l_z}$ 分布 (b)

3.2 三维迂回相位编码：对正空间非线性结构多值化编码

在上一节通过三维 GS 算法算出了正空间中 $16 \times 16 \times 16$ 个 $\Lambda_x \times \Lambda_y \times \Lambda_z = 4 \times 3 \times 4 \mu\text{m}$ 小长方体中每个长方体内的相位分布后，倒格矢 $\mathbf{G}_{2,1,0}$ 所指向的倒格子内， $16 \times 16 \times 16$ 个细节倒格矢 $(q_{l_x}, q_{l_y}, q_{l_z})$ 所对应的细节倒格子的角谱分布 $C_{2,1,0;l_x,l_y,l_z}$ 的振幅分布，也就大致是图 3.4 的目标振幅分布了；可见在 16×16 的低像素的情况下，至少在轮廓即 0-1 分布上是完全一致的，只不过振幅大小略有不同。因此振幅正比于该倒空间中三维光子晶体结构的角谱的振幅分布的倍频光集 $\{E_{3;m_x,m_y,m_z;l_x,l_y,l_z}\}$ ，经过夫琅禾费衍射后在远场所成的正空间的像，便是目标图案的光强分布。

接下来要做的便是将三维 GS 算法所得的正空间中 $16 \times 16 \times 16$ 个 $4 \times 3 \times 4 \mu\text{m}$ 的小长方体，所对应的 $16 \times 16 \times 16$ 大小的三维相位分布矩阵，通过逐层二维迂回相位编码的方法，将三维矩阵中每一个像素所储存相位信息，落实到相应正空间坐标处的周期性小长方体结构中去。

总调制区域即尺寸为 $64 \times 48 \times 64 \mu\text{m}$ 的正空间大长方体在 y 方向上的厚度为 48 μm ，若以 $1 \mu\text{m}$ 为 1 层则共有 48 层。但由于飞秒激光直写头的移动精度，允许我们再将这 y 向的 48 层进一步划分为 480 小层，每一层 $0.1 \mu\text{m}$ ，10 小层为 1 大层，3 大层为 1 个小长方体在 y 方向的周期 $\Lambda_y = 3 \mu\text{m}$ ，因此每一个小长方体在这第二部分的仿真程序中，被划分为了 30 个 y 向厚度为的 $0.1 \mu\text{m}$ 的小层，注意到该厚度比匹配该级别的光波波长 $0.859 \mu\text{m}$ 要小大约 1 个数量级，因此调制出来的三维光子晶体的二

阶非线性系数在实空间中的三维分布精度对于实现三维傅里叶全息成像是够高的。

针对这 480 层在正空间调制区域中以 $x-o-z$ 向二维分布的 16×16 个 $4 \times 4 \mu\text{m}$ 的正方形小方块，需逐层采用二维迂回相位编码方法将三维 GS 算法所得的正空间相位分布转换为三维光子晶体调制区域中每一层的实际结构。

3.2.1 二维迂回相位编码原理

3.2.1.1 总可用不同复振幅分布的傅里叶全息图重现有限范围内的同一个目标光场

二维迂回相位编码技术最早被应用于二元振幅型傅里叶迂回计算全息图的生成。

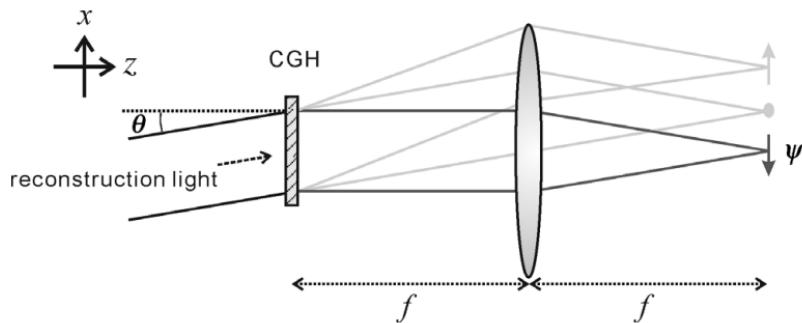


图 3.5 二维傅里叶计算全息还原重现像的过程

在典型的傅里叶计算全息图还原重建像的过程中，若不考虑零级背景光和孪生像的问题，则二维傅里叶全息图的复振幅 $H(x, y)$ 与二维重建像的复振幅 $\psi(x, y)$ 满足：

$$\psi(x, y) = \mathcal{F}\{H(x, y)\psi_r(x, y)\} \quad (3.4)$$

其中， $\psi_r(x, y) = e^{ik_3 x} = e^{ik_3 \cdot \sin \theta \cdot x} = e^{ik_3 \cdot x_0 / f \cdot x}$ 为不考虑振幅的斜射单色平面波重现光 (reconstruction light)，因它的斜射而在 CGH(Computer-generated hologram)、透镜、光屏三个互相平行的平面上，引入了相同的附加相位因子，且其本应是 (x, y) 的函数，但因参考光垂直于 y 轴沿纸面入射而使得其在引入的额外相位与 y 无关。

这里把透过 CGH 前后的重现光比作三维光子晶体中经过 $64 \times 48 \times 64 \mu\text{m}$ 的调制区域前后的倍频光 E_3 ，相应的傅里叶计算全息图 CGH 比作三维光子晶体中的二阶非线性系数的调制区域，CGH 中储存的信息等价于调制区域的 $\chi_{eff}^{(2)}(\mathbf{r})$ 的调制函数分布信息即光子晶体中所刻写的结构信息，携带 CGH 信息的重现光经过透镜比作携带光子晶体结构信息的倍频光 E_3 经过远场夫琅禾费衍射，重现光在光屏上成的像 $\psi(x, y)$ 比作倍频光 E_3 的远场像，所以重现光沿用了 E_3 的波矢大小 k_3 。

针对光屏上有限范围 $d_x \times d_y$ 内的目标光场 $\psi(x, y)$ ，二维迂回相位编码所生成的二元(binary)振幅型 CGH (开孔或不开孔对应透光或不透光) 即 $H_b(x, y)$ ，可以替代实际

的傅里叶全息图的复振幅分布 $H(x, y)$ ，经傅里叶变换后复现相同的重建像：

$$\mathcal{F}\{H_b(x, y)\psi_r(x, y)\} = \mathcal{F}\{H(x, y)\psi_r(x, y)\} = \psi(x, y) \quad (3.5)$$

3.2.1.2 原复振幅傅里叶计算全息图的采样全息图及其相应的采样重建光场分布

将 $H(x, y)$ 划分为 $M \times N$ ($M = N$) 个 $w \times w$ 的格子（虚拟网格），取其中某个正方形小格子中心为坐标原点，其余小格子的中心位置便可由 $(x_m, y_n) = (m \cdot w, n \cdot w)$ 表示，其中 (m, n) 的取值范围由 (M, N) 及其奇偶性确定：

$$\begin{cases} m_{\min} = -\frac{M - \text{mod}(M, 2)}{2}, m_{\max} = -\frac{M - \text{mod}(M, 2)}{2} + M - 1 = \frac{M + \text{mod}(M, 2)}{2} - 1 \\ n_{\min} = -\frac{N - \text{mod}(N, 2)}{2}, n_{\max} = \frac{N + \text{mod}(N, 2)}{2} - 1 \end{cases} \quad (3.6)$$

对于三维光子晶体调制范围 $64 \times 48 \times 64 \mu\text{m}$ 的每一个 $x-o-z$ 面，由于其由 16×16 个 $4 \times 4 \mu\text{m}$ 的小正方形组成，所以可认为实际情况下，那里的 $M = N = 16$ 、 $w = 4 \mu\text{m}$ 。如下图所示：

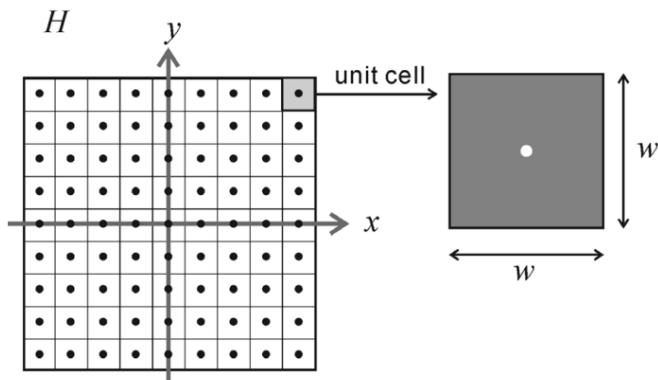


图 3.6 复振幅傅里叶全息图的小格子划分

对于任意一个尺寸 $(M \cdot w) \times (N \cdot w)$ 已经确定了的 $H(x, y)$ ，总可以将 M, N 调整得够大，或者将小格子边长 w 划分得足够小、小格子足够密集，使得 $H(x, y) = A(x, y) \times e^{i\phi(x, y)}$ 在每个小方格中的值的起伏不大，此时对 $H(x, y)$ 在每个小方格中心处的振幅和相位即 $A(x_m, y_n)$ 与 $e^{i\phi(x_m, y_n)}$ 进行采样，其值近似于整个小方格区域内 $A(x, y)$ 与 $e^{i\phi(x, y)}$ 的平均值，同一小方格内其他 (x, y) 处的 $A(x, y)$ 与 $e^{i\phi(x, y)}$ 与其相差不大，因此对 $H(x, y)$ 可以用格点个数个离散的点光源阵列近似表示：

$$\begin{aligned} H(x, y) &\approx H_s(x, y) = \sum_{m, n} A(x_m, y_n) \times e^{i\phi(x_m, y_n)} \cdot \text{rect}\left(\frac{x - x_m}{w}, \frac{y - y_n}{w}\right) \\ H_s(x, y) &= \sum_{m, n} A_{mn} \times e^{i\phi_{mn}} \cdot \text{rect}\left(\frac{x - x_m}{w}, \frac{y - y_n}{w}\right) \end{aligned} \quad (3.7)$$

其中 s 表示 sampled, 当 w 足够小时, $\text{rect}[(x - x_m)/w, (y - y_n)/w]$ 函数变为 Kronecker delta 即 $\delta_K[x - x_m, y - y_n]$ 而非 $\delta_D[x - x_m, y - y_n] = \text{rect}[(x - x_m)/w, (y - y_n)/w]/w$ 即 Dirac delta。

对于 $H_s(x, y)$, 此时重现像的光场分布也是采样式的:

$$\begin{aligned}
\psi_s(x, y) &= \mathcal{F}\{H_s(x, y)\psi_r(x, y)\} = \sum_{m,n} A_{mn} \times e^{i\phi_{mn}} \cdot \mathcal{F}\left\{\text{rect}\left(\frac{x - x_m}{w}, \frac{y - y_n}{w}\right) e^{ik_{3x}x}\right\} \\
&\xrightarrow{\text{位移定理}} \sum_{m,n} A_{mn} \times e^{i\phi_{mn}} \cdot \frac{1}{4\pi^2} \left\langle e^{i[(k_x - k_{3x})x + k_y y]} \middle| \text{rect}\left(\frac{x - x_m}{w}, \frac{y - y_n}{w}\right) \right\rangle \\
&\xrightarrow{\text{延迟定理}} \sum_{m,n} A_{mn} \times e^{i\phi_{mn}} \cdot \frac{1}{4\pi^2} e^{-i[(k_x - k_{3x})x_m + k_y y_n]} \left\langle e^{i[(k_x - k_{3x})x + k_y y]} \middle| \text{rect}\left(\frac{x}{w}, \frac{y}{w}\right) \right\rangle \\
&\xrightarrow{\text{相似性定理}} \sum_{m,n} A_{mn} \times e^{i\phi_{mn}} \cdot \frac{w^2}{4\pi^2} e^{-i[(k_x - k_{3x})x_m + k_y y_n]} \left\langle e^{iw[(k_x - k_{3x})x + k_y y]} \middle| \text{rect}(x, y) \right\rangle \\
&= \frac{w^2}{4\pi^2} \sum_{m,n} A_{mn} \times e^{i\phi_{mn}} \cdot e^{ik_{3x}x_m} e^{-i[k_x x_m + k_y y_n]} \frac{\sin \frac{w(k_x - k_{3x})}{2}}{\frac{w(k_x - k_{3x})}{2}} \frac{\sin \frac{wk_y}{2}}{\frac{wk_y}{2}} \\
&= \frac{w^2}{4\pi^2} \sum_{m,n} A_{mn} \times e^{i\phi_{mn}} \cdot e^{ik_{3x}x_m} e^{-i[k_x x_m + k_y y_n]} \text{sinc}\left[\frac{w(k_x - k_{3x})}{2}, \frac{wk_y}{2}\right]
\end{aligned} \tag{3.8}$$

其中 $k_x = k_0 \cdot \sin \theta_x = k_0 \cdot x/f$ 、 $k_y = k_0 \cdot \sin \theta_y = k_0 \cdot y/f$ 类似 $\chi_{\text{eff}}^{(2)}(\mathbf{r})$ 所对应的 \mathbf{G}_{m_x} 、 \mathbf{G}_{m_y} , 均属于描述正空间结构的频谱信息的频谱空间中的坐标量, 与光波长无关 ($k_0 \neq k_3$)。即使 w 相对于 d_x 并不足够小, 只要是远场成像即 $d_x/f \ll 1$, 且 k_0 不大, 则有:

$$\left\{ \begin{array}{l} \left| \frac{w(k_x - k_{3x})}{2} \right|_{\max} = \left| \frac{w(k_0 x - k_3 x_0)}{2f} \right|_{\max} = \left| \frac{w(k_0 d_x/2 - k_3 x_0)}{2f} \right| \ll 1 \\ \left| \frac{wk_y}{2} \right|_{\max} = \frac{wk_0 d_y}{2f} \ll 1 \end{array} \right. \tag{3.9}$$

在该条件下, 上式(3.8)中的二维 sinc 函数的值趋近 1, 此时该式可理解为对原 $H(x, y)$ 的采样全息图 $H_s(x, y)$ 上, 无数个点光源所发出的平面波, 在重现平面的叠加。当进一步设定小方格尺寸为:

$$w = \frac{2\pi}{k_{3x}} l = \frac{2\pi l}{k_3 \cdot \sin \theta} = \frac{l \cdot \lambda_3}{\sin \theta}, \quad (l \in \mathbb{N}^+) \tag{3.10}$$

时, 重现光所引入的斜射因子 $e^{ik_{3x}x_m} = e^{ik_3 \cdot \sin \theta \cdot m \cdot w} = e^{i2\pi l \cdot m} = 1$ 将消失, 此时重建像的采样光场分布为:

$$\psi_s(x, y) \approx \frac{w^2}{4\pi^2} \sum_{m,n} A_{mn} \times e^{i\phi_{mn}} \cdot e^{-i[k_x x_m + k_y y_n]} \tag{3.11}$$

将其与式(3.7)对比, 可见二者拥有相同的采样部分 $\sum_{m,n} A_{mn} \times e^{i\phi_{mn}}$ 。

在三维非线性全息中, 当周期性平面结构的尺寸 $w = 4 \mu\text{m}$ 和重现光即倍频光波长 $\lambda_3 = 0.859 \mu\text{m}$ (3.3 节会进一步修订为 $0.85 \mu\text{m}$) 给定后, 由于重现光即倍频光在正空

间中是正入射调制区域，因此式(3.10)中的 l 不论如何取值，哪怕不取正整数而取分数或无理数，此时正入射所引入的额外相位因子均因 $\sin \theta = 0$ 而为1。因此在三维非线性全息的正入射情况下， l 可以任意取值，且 l 与 $\sin \theta$ 的值均与倒格矢 $\mathbf{G}_{2,1,0}$ 的方位无关，波长 λ_3 与 l 也不必满足式(3.10)的关系，因其右侧分母 $\sin \theta = 0$ 。

3.2.1.3 一种 $H_b(x, y)$ 即二元振幅型傅里叶全息图的开孔策略

基于对 $H(x, y)$ 划分完了虚拟网格后的采样傅里叶全息图 $H_s(x, y)$ ，现取另一形状和尺寸均与 $H_s(x, y)$ 相同的不透光光阑，并也划分好与之相同尺寸的虚拟网格即 $M \times N$ ($M = N$)个 $w \times w$ ，在其每一个小方格中，以其中心 $(x_m, y_n) = (m \cdot w, n \cdot w)$ 为小方格坐标系原点，对每一个小方格中的不透光材料进行开孔，接下来将推出开孔面积 $a_{mn} = p_{mn} \times q_{mn}$ 正比于透光振幅 A_{mn} ，则可通过调整每一个孔的面积以控制该小方格所代表的点光源光场振幅，进而调控整个二维全息面的光场振幅分布 $A(x, y)$ 、 $A_{mn} = A(x_m, y_n)$ ；而调整每个矩形孔中心相对于小方格中心的(x 向)位移 α_{mn} ，即可单独改变每个小方格内次波源的初始相位，进而控制 $\phi(x, y)$ 、 $\phi_{mn} = \phi(x_m, y_n)$ 即 $H_b(x, y)$ 的二维相位分布。

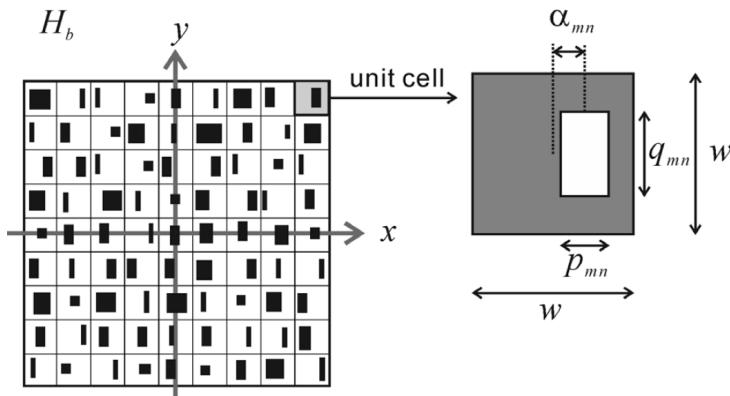


图 3.7 迂回相位二元振幅全息图示意图

对于二元振幅型傅里叶全息图 $H_b(x, y)$ ，即阵列式开孔的光阑，其透过率函数即为其复振幅分布，同时也为其振幅分布，因其相位分布是均匀的：

$$H_b(x, y) = \sum_{m,n} \text{rect}\left(\frac{x - x_m - \alpha_{mn}}{p_{mn}}, \frac{y - y_n}{q_{mn}}\right) \quad (3.12)$$

在相同的斜射等振幅单色平面重现光 $\psi_r(x, y) = e^{ik_3 x}$ 的照射下，重建像所在的平面的复振幅分布为：

$$\psi_b(x, y) = \mathcal{F}\{H_b(x, y)\psi_r(x, y)\} = \sum_{m,n} \mathcal{F}\left\{\text{rect}\left(\frac{x - x_m - \alpha_{mn}}{p_{mn}}, \frac{y - y_n}{q_{mn}}\right)e^{ik_3 x}\right\}$$

$$\begin{aligned}
& \xrightarrow{\text{位移定理}} \sum_{m,n} \frac{1}{4\pi^2} \left\langle e^{i[(k_x - k_{3x})x + k_y y]} \left| \text{rect}\left(\frac{x - x_m - \alpha_{mn}}{p_{mn}}, \frac{y - y_n}{q_{mn}}\right) \right. \right\rangle \\
& \xrightarrow{\text{延迟定理}} \sum_{m,n} \frac{1}{4\pi^2} e^{-i[(k_x - k_{3x})(x_m + \alpha_{mn}) + k_y y_n]} \left\langle e^{i[(k_x - k_{3x})x + k_y y]} \left| \text{rect}\left(\frac{x}{p_{mn}}, \frac{y}{q_{mn}}\right) \right. \right\rangle \\
& \xrightarrow{\text{相似性定理}} \sum_{m,n} \frac{p_{mn} q_{mn}}{4\pi^2} e^{-i[(k_x - k_{3x})(x_m + \alpha_{mn}) + k_y y_n]} \left\langle e^{i[p_{mn}(k_x - k_{3x})x + q_{mn}k_y y]} \left| \text{rect}(x, y) \right. \right\rangle \\
& = \frac{1}{4\pi^2} \sum_{m,n} p_{mn} q_{mn} e^{ik_{3x}x_m} e^{-i(k_x - k_{3x})\alpha_{mn}} e^{-i[k_x x_m + k_y y_n]} \frac{\sin \frac{p_{mn}(k_x - k_{3x})}{2}}{\frac{p_{mn}(k_x - k_{3x})}{2}} \frac{\sin \frac{q_{mn}k_y}{2}}{\frac{q_{mn}k_y}{2}} \\
& = \frac{1}{4\pi^2} \sum_{m,n} p_{mn} q_{mn} e^{ik_{3x}x_m} e^{-i(k_x - k_{3x})\alpha_{mn}} e^{-i[k_x x_m + k_y y_n]} \text{sinc}\left[\frac{p_{mn}(k_x - k_{3x})}{2}, \frac{q_{mn}k_y}{2}\right] \quad (3.13)
\end{aligned}$$

由于在虚拟网格的每个小方格内所开孔径必小于小方格边长即 $p_{mn}, q_{mn} < w$ ，则根据式(3.9)的理由，上式中 sinc 函数值趋近于 1；再依式(3.10)对光阑的虚拟网格尺寸设置与 $H_b(x, y)$ 的网格相同的边长 $w_l = 2\pi l/k_{3x}$ ，则上式中倾斜因子 $e^{ik_{3x}x_m} = 1$ 也可略去；又因 $d_x/x_0 \ll 1$ ，而有

$$\frac{|k_x|_{\max}}{k_{3x}} = \frac{(k_0 d_x/2)/f}{k_3 x_0/f} = \frac{k_0 d_x/2}{k_3 x_0} \ll 1 \quad (3.14)$$

所以式(3.13)中的 $e^{-i(k_x - k_{3x})\alpha_{mn}} \approx e^{ik_{3x}\alpha_{mn}}$ ，于是代入开孔面积 $a_{mn} = p_{mn}q_{mn}$ 有：

$$\psi_b(x, y) \approx \frac{1}{4\pi^2} \sum_{m,n} a_{mn} e^{ik_{3x}\alpha_{mn}} e^{-i[k_x x_m + k_y y_n]} \quad (3.15)$$

对比式(3.11)得：

$$\begin{cases} a_{mn} = w^2 A_{mn} \\ k_{3x} \alpha_{mn} = \phi_{mn} \end{cases} \quad (3.16)$$

代入 $w = 2\pi l/k_{3x}$ 得：

$$\begin{cases} a_{mn} = w^2 A_{mn} \propto A_{mn} \\ \alpha_{mn} = \frac{\phi_{mn}}{k_{3x}} = \frac{\phi_{mn}}{2\pi} \frac{w}{l} \propto \phi_{mn}, \alpha_{mn} \in \frac{[-\pi, \pi]}{2\pi} \frac{w}{l} = [-\frac{w}{2l}, \frac{w}{2l}] \end{cases}, (l \in \mathbb{N}^+) \quad (3.17)$$

其中孔的位移量 α_{mn} 不仅受到式(3.17)的限制，还要用 l 的取值，或 p_{mn} 配合 ϕ_{mn} 的取值，或 ϕ_{mn} 的取值范围，来保证小孔边界没有超出虚拟周期性小方格的边界，即

$$|\alpha_{mn}| + \frac{p_{mn}}{2} \leq \frac{w}{2} \Rightarrow |\alpha_{mn}| \leq \frac{w - p_{mn}}{2} \quad (3.18)$$

在实际中采用设置 $l = 2$ ，以及 $p_{mn} = 1.5 \mu\text{m}$ 来保证上式的成立。

而孔的面积 a_{mn} 暗示光阑上所开的孔不需要是严格的矩形孔，可以是圆孔、椭圆孔甚至任何其他形状，但其面积必须是 a_{mn} ，同时其中心相对于小方格的中心的位移也必须是上式所确定的 α_{mn} 。该式便给了我们对不透光正方形光阑设计每个开孔的位移和开孔面积的依据。

3.2.2 三维迂回相位编码过程

三维光子晶体的逐层编码便采用了二维迂回相位开孔策略，可认为其正空间中 480 层 $0.1 \mu\text{m}$ 厚的 $x-o-z$ 向 16×16 个 $4 \times 4 \mu\text{m}$ 的正方形小方块，分别对应了 $M = N = 16$ 、 $w = 4 \mu\text{m}$ 、 $d_x = d_y = 64 \mu\text{m}$ ，每一层的相位分布，即式(3.17)或式(3.7)中的 ϕ_{mn} ，也即生成 3.1.4.2 节中图 3.4 中右侧 16 张 16×16 像素的角谱振幅分布图的，下面的 16 张 16×16 像素的纯相位分布图中，某张图的二维相位分布：

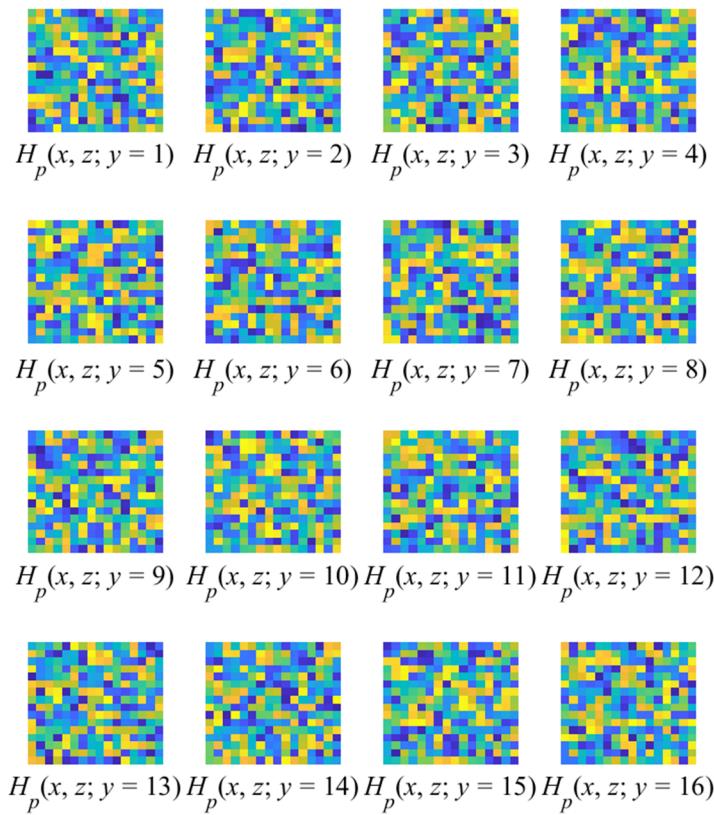


图 3.8 三维 GS 算法所生成的正空间二阶非线性系数的三维相位分布

在 480 层待编码的平面中，处于同一个 $\Lambda_y = 3 \mu\text{m}$ 内的 30 层厚 $0.1 \mu\text{m}$ 的待编码平面，均共用同一个 $x-o-z$ 面的二维相位分布，即共用图 3.8 中某一张 16×16 像素的图片。但不能 30 层中每一层都按照相应的二维相位分布进行编码和加工，因为如果这样做，则 y 向的结构周期便不再是 $\Lambda_y = 3 \mu\text{m}$ ，而是对于各 (x, z) 没有统一的周期可言。受到一维准相位匹配结构在一个周期内的正负畴宽度相等时基频光到倍频光的转

换效率最高的启发， y 向一个 $\Lambda_y = 3 \mu\text{m}$ 周期内的编码与未编码、加工与未加工的层状结构宽度比应为1:1，即前 $1.5 \mu\text{m}$ 内的15层编码，后 $1.5 \mu\text{m}$ 内的15层不编码，或反之。否则倍频效率会降低，以致最终全息成像图案整体亮度下降。仿真也证实了当编码占空比为1:4和1:1时，下一节模拟的远场成像整体光强没有占空比为1:2时强。

对于 y 向一个周期内，前半个周期中的每一层，由于其不是不透光的光阑因而对其进行的迂回相位编码操作，不是对其内每个 4×4 的小方格进行开孔，而是对其进行飞秒激光直写，本身应开孔的位置现经飞秒激光短暂作用后，相应位置区域的二阶非线性系数 $\chi_{eff}^{(2)}(\mathbf{r})$ 的值会降低15%，因此式(3.17)中的 a_{mn} 在三维非线性全息中的含义并非开孔面积，而是激光光斑在 $x-o-z$ 面的烧灼面积，由于已经证实开孔形状不一定需要为矩形，因此即使激光光斑在 $x-o-z$ 方向的横截面是圆形或椭圆形，也不影响式(3.17)中相应结论的成立。而激光打击位置即光斑中心相对于 $x-o-z$ 向平面周期结构即小正方形中心的 x 向偏移量 α_{mn} ，符合式(3.17)的约束。

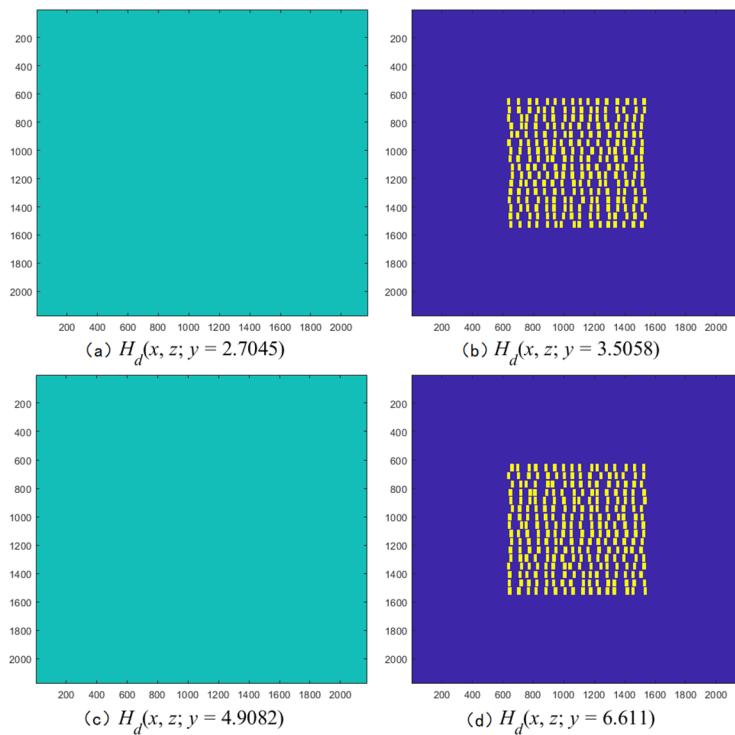


图 3.9 三维迂回相位编码的各层二维开孔结构（左列相对右列为同一周期内的非编码面）

我们的设备所产生的飞秒激光光斑直径 $\sim 1.2 \mu\text{m}$ ，纵深 $5 \sim 10 \mu\text{m}$ ，因此设定了 $p_{mn} = 3.5 \mu\text{m}$ 、 $q_{mn} = 1.5 \mu\text{m}$ 的矩形孔为编码单元，面积与激光光斑纵截面面积相仿，且宽度较窄，满足式(3.18)，以至小孔虽在 x 方向有位移，但总会落在 4×4 小方格中。

由于加工系统的平移理论精度为 10 nm ，因此不仅即将 $\Lambda_y = 3 \mu\text{m}$ 分成了30层，还在 x, z 向上将 $\Lambda_x = \Lambda_z = 4 \mu\text{m}$ 分别分成了60层，调制区域在 x, z 方向一共被划分为了

960 层（见下图光斑分布区域的横纵向尺寸），因此正空间中 $960 \times 480 \times 960$ 个最小移动步长/增量所构成的密集细网格，能够保证孔中心的实际位移量 α_{mn} 误差较小，其所代表的相位值就较精确，且位移和相位可近似连续取值。由于仿真中设置了 $l=2$ 使得 $\alpha_{mn} \in [-w/2l, w/2l] = [-w/4, w/4] = [-1, 1] \mu\text{m}$ ，而在该区间内有 30 层，因此对位移量 α_{mn} 和相位 ϕ_{mn} 的分级为 30 级，即 $[-15, 15]/15 \times \mu\text{m} \leftrightarrow [-\pi, \pi]$ 。

上述将小孔面积取值与光斑面积相等，并基于此以确定方形小格子尺寸 $w \times w$ 的设计，是为了在加工空间范围 $d_x \times d_y$ 固定的条件下，实现最大的小格子数 $M \times N$ ，这样目标图像的像素和容量、信息便可更大，目前因加工所用飞秒激光光斑面积无法再小， 16×16 像素的目标图片已经是最大尺寸了。并且由于飞秒激光光斑既不能更小，在确定了最小 $w \times w$ 尺寸后光斑又不能更大，所以光斑即小孔的面积无法改变，则式(3.17)中的式一即 a_{mn} 就无法改变，以至于振幅分布 A_{mn} 就无法改变，这就要求必须使用三维 GS 算法得出傅里叶全息图的纯相位分布，并只能用上式(3.17)的式二，即孔位移量分布来代表相位分布，无法利用和表示傅里叶全息图中的振幅分布信息。

若将开孔面积取得比光斑面积大，则虚拟小方格尺寸 $w \times w$ 也会增大，导致像素 $M \times N$ 降低，允许加载的目标图片像素就会降低；但这样做有个好处，即小孔面积所对应的大光斑面积，将由许多小光斑面积的交集所得，而不是一个小光斑的直接结果，是积分效应；因此此时不仅小孔形状所对应的大光斑形状，可与矩形差别不大、可以是任意形状，而且可以通过调控小孔即大光斑大小，以调控傅里叶全息图振幅分布，这样便不再需要通过三维 GS 算法，计算等效于原傅里叶全息图的三维复振幅分布的纯相位全息图的三维相位分布。

3.3 三维光子晶体各面倍频光光强、相位分布及远场成像仿真

3.1 节通过三维 GS 算法完成了从目标二维图像到实空间中晶体调制区域内三维相位分布的转化；3.2 节利用三维迂回相位编码法将 3.1 节所得的三维相位分布转换为了三维光子晶体调制区域内的实际刻写的结构分布；该 3.3 节将验证 3.2 节所编码和刻写完毕的三维光子晶体，其结构是否能实现基频光到倍频光的转换的同时，倍频光在远场的成像的特定衍射级次的光场分布，也为 3.1 节的目标图像。

在仿真程序的第三部分中，将第二部分所生成的 480 层二维结构分布，所对应的 480 个 mat 文件，以及 1 个储存其他参数信息的 mat 文件，作为第三个程序，即三维光子晶体内光强分布，以及远场光强分布模拟程序的输入，所得的输出即基频光场在

晶体内部既实现倍频的同时，又被所经过的 $64 \times 48 \times 64 \mu\text{m}$ 的调制区域调制了其倒空间二维频谱信息后的倍频光场，出射晶体并经夫琅禾费衍射后所得的远场光强分布。

3.3.1 三维光子晶体两侧面的倍频光强分布、相位分布

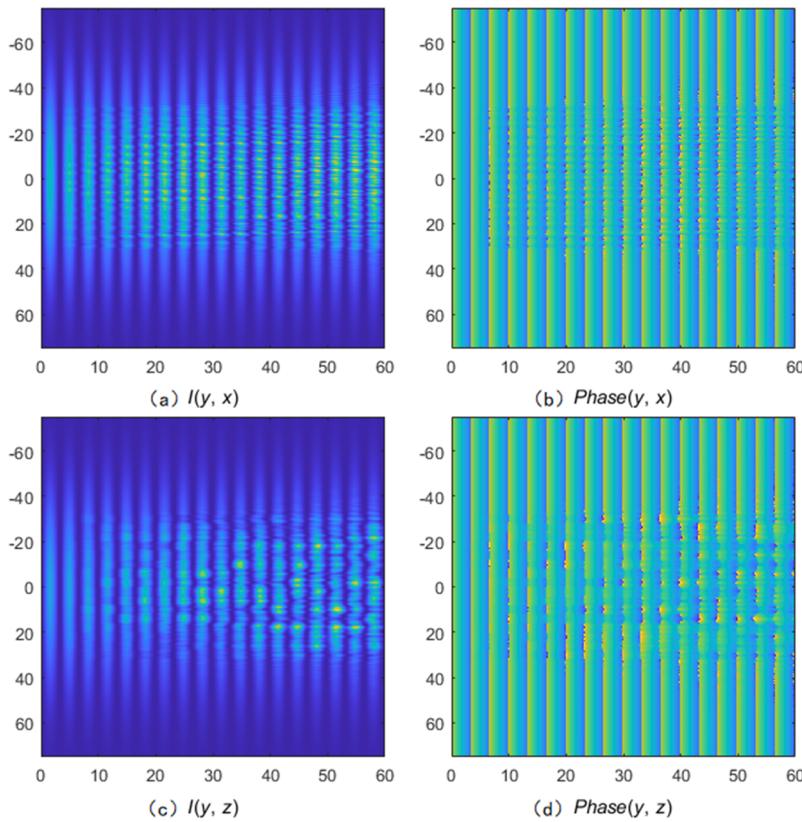


图 3.10 晶体中轴面 $y-o-x$ 面（上排）和 $y-o-z$ 面（下排）的光强分布和相位分布

图 3.10 和图 3.11 的横纵坐标单位均为 μm ，图 3.10 中带光斑点的区域在通光方向的调制区域厚度，对应横轴即 y 向的 $48 \mu\text{m}$ ，调制区域在纵轴即 x 向和 z 向的宽度和高度为 $64 \mu\text{m}$ ，除此之外 y 向还空余了 $12 \mu\text{m}$ ， x 向和 z 向也分别剩下 $86 \mu\text{m}$ ，是晶体内部未调制区域；调制区域与视野区域的尺寸占比 $64:150 = 960:2200$ 与图 3.9 相同。

以上四张图的 y 向光强和相位起伏周期相同且大致为 $3.33 \mu\text{m}$ ，这可通过查看式 2.115 或 2.116 在 $\hat{\mathbf{L}}_{k_3}, k_3$ 指向 y 向时的周期得出：该方向调制区域的光强起伏周期与非调制区域的大致相同，毕竟所选级次的倒格矢 $\mathbf{G}_{2,1,0}$ 与细节倒格矢 $\mathbf{q}_{l_x,6,l_z}$ 的大小，相对于波矢失配量 $\Delta\mathbf{k}$ 而言比较小，所以不用考察 $\Delta\mathbf{k}_o$ 对应的调制区域周期，而是直接考察 $\Delta\mathbf{k}$ 所对应的非调制区域的光强起伏周期，可见理论计算结果与实际仿真结果吻合：

$$\frac{\Delta\mathbf{k} \cdot \hat{\mathbf{L}}_{k_3}}{2} = \pi \Rightarrow L_{k_3} = \frac{2\pi}{|\Delta\mathbf{k}_y|} = \frac{2\pi}{k_{3y} - k_{1y} - k_{2y}} = \frac{2\pi}{\frac{2\pi}{0.85/2/2.2909} - 2 \times \frac{2\pi}{0.85/2.1628}} = 3.318 \mu\text{m} \quad (3.19)$$

3.3.2 三维光子晶体后端面与远场倍频光强分布

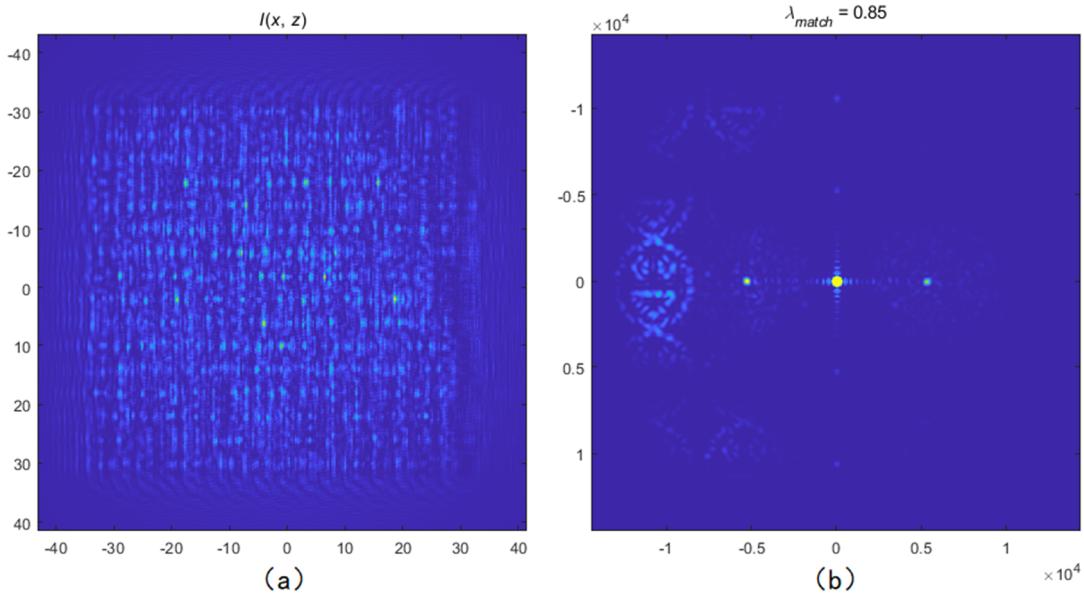


图 3.11 晶体后端面 $x-o-z$ 面光强分布 (a) 及 $G_{2,1,0}$ 所对应的倍频光远场二级衍射成像 (b)

选择不同的基波波长，可匹配到不同的衍射级 m_x, m_y, m_z 或倒格矢 $\mathbf{G}_{m_x, m_y, m_z}$ 所对应的笑脸图案。此前在 3.1 节中暂时确定的 $0.859 \mu\text{m}$ 的基波波长，不一定准确匹配 $G_{2,1,0}$ 该级倒格矢，需在 $0.859 \mu\text{m}$ 附近改变波长查看远场成像中相应级次的目标图案是否是中部较亮，才能确定最终匹配该级倒格矢或衍射级次的基波波长，最终仿真得到的匹配波长为 $0.850 \mu\text{m}$ 。

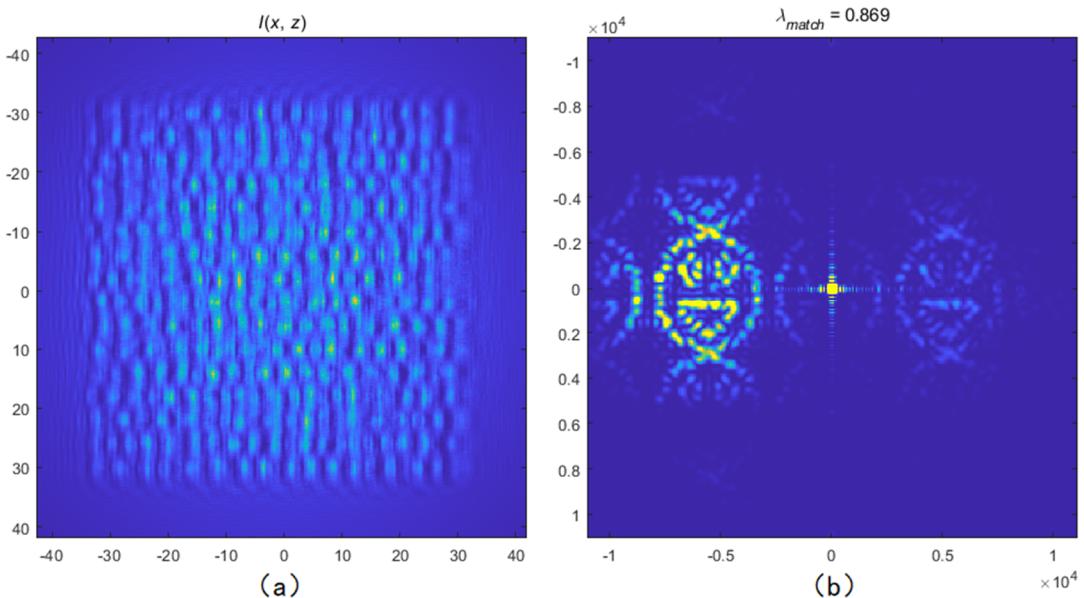


图 3.12 $G_{1,1,0}$ 所对应的倍频光远场一级衍射成像、晶体后端面光强分布

之所以改变波长可以匹配不同的衍射级次，且不同衍射级的图案对应不同的远场

x 向位置，是因不同基波波长的晶体内倍频波矢曲面大小不同，导致波矢失配量不同，以至于起点固定在两倍晶体内基波波矢处、终点落在晶体内倍频光波矢曲面上的倒格矢 $\mathbf{G}_{m_x, m_y, m_z}$ 不同，则对应不同细节倒格矢 $\mathbf{q}_{l_x, l_y, l_z}$ 的、大小与方向均在 $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{G}_{m_x, m_y, m_z}$ 即 $\mathbf{k}_{3; m_x, m_y, m_z}$ 附近的波矢集 $\{\mathbf{k}_{3; m_x, m_y, m_z; l_x, l_y, l_z}\} = \{\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{G}_{m_x, m_y, m_z} + \mathbf{q}_{l_x, l_y, l_z}\}$ 的公有出射方向，便因 $\mathbf{G}_{m_x, m_y, m_z} = 2\pi m_x / \Lambda_x \mathbf{i} + 2\pi m_y / \Lambda_y \mathbf{j} + 2\pi m_z / \Lambda_z \mathbf{k}$ 即 m_x, m_y, m_z 的不同而不同，且 $0 < m_x = 1 < 2$ 对应的 $\mathbf{k}_{3x; 1, 1, 0} = \mathbf{G}_{m_x} \Big|_{m_x=1} = 2\pi / \Lambda_x$ ，便会比 $m_x = 2$ 对应的 $\mathbf{k}_{3x; 2, 1, 0} = \mathbf{G}_{m_x} \Big|_{m_x=2} = 4\pi / \Lambda_x$ 偏 x 轴负向即偏右^①，同样比 $m_x = 0$ 对应的 $\mathbf{k}_{3x; 0, 1, 0} = 0$ 偏左，这便解释了图 3.12 中拥有 $\mathbf{k}_{3; 1, 1, 0}$ 邻域内的波矢集 $\{\mathbf{k}_{3; 1, 1, 0; l_x, l_y, l_z}\}$ 的簇射倍频光集 $\{E_{3; 1, 1, 0; l_x, l_y, l_z}\}$ 所生成的一级衍射目标笑脸图，为何在二级衍射笑脸图像 $\{E_{3; 2, 1, 0; l_x, l_y, l_z}\}$ 的右侧，同时又在零级衍射图像的左侧。

而之所以一级衍射成像整体背景强度比二级衍射成像高，可以这样解释：据式(2.117)有 $|E_{3; m_x, m_y, m_z; l_x, l_y, l_z}| \propto C_{m_x, m_y, m_z; l_x, l_y, l_z}$ ，而据式(2.101)有：

$$C_{m_x, m_y, m_z; l_x, l_y, l_z} = \frac{1}{\Omega} \iiint_{\Omega} B_{\Omega}(\mathbf{r}) e^{-i(\mathbf{G}_{m_x, m_y, m_z} + \mathbf{q}_{l_x, l_y, l_z}) \cdot \mathbf{r}} d^3 r \approx \frac{1}{\Omega} \iiint_{\Omega} \frac{\chi_{eff; \Omega}^{(2)}(\mathbf{r})}{\chi_{eff}^{(2)}} e^{-i(\mathbf{G}_{m_x, m_y, m_z} + \mathbf{q}_{l_x, l_y, l_z}) \cdot \mathbf{r}} d^3 r \quad (3.20)$$

则将三维 GS 算法所得的正空间相位分布代入上式会发现（从一维准相位匹配周期性结构的角谱分布图 2.8 可见一斑）：

$$C_{1, 1, 0; l_x, l_y, l_z} > C_{2, 1, 0; l_x, l_y, l_z} \quad (3.21)$$

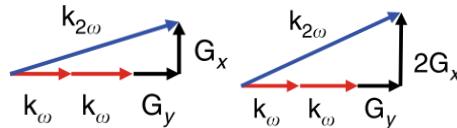


图 3.13 $\{C_{1, 1, 0; l_x, l_y, l_z}\}$ 与 $\{C_{2, 1, 0; l_x, l_y, l_z}\}$ 所对应的 $\{\mathbf{k}_{3; 1, 1, 0; l_x, l_y, l_z}\}$ 与 $\{\mathbf{k}_{3; 2, 1, 0; l_x, l_y, l_z}\}$ 的大致方位^[14]

即集合 $\{C_{1, 1, 0; l_x, l_y, l_z}\}$ 内每一个元素的值都比集合 $\{C_{2, 1, 0; l_x, l_y, l_z}\}$ 内具有相同细节倒格矢元素的值更大。于是便有 $|E_{3; 1, 1, 0; l_x, l_y, l_z}| > |E_{3; 2, 1, 0; l_x, l_y, l_z}|$ 。这样一来，图 3.11 的左图即携带结构倒空间一级衍射信息的正空间光强分布，便比图 3.12 的左图携带倒空间二级衍射信息的整体光强分布更强，再根据帕色渥定理或能量守恒定理， $m_x = 1$ 级与 $m_x = 2$ 级的晶体后端面光强分布，经过夫琅禾费衍射即一次傅里叶变换前后，二者在频域中的

^① 入射光入射方向和 y 轴负向均朝右，且从右往左看晶体后端面的 x -o-z 系的 x 轴朝右，则 x 轴插入纸面向里， m_x, m_y, m_z 、 $\mathbf{G}_{m_x, m_y, m_z}$ 与 $\mathbf{k}_{3; m_x, m_y, m_z}$ 均参照该坐标系。而对于最终出的图像，其所参照的坐标系为 x 轴朝右，对应穿出纸面向外， z 轴朝上不变。

能量与其在空域中的能量相同，所以最终正空间中光强分布是一级衍射图像大于二级衍射图像。

还有一种不可忽略的因素会直接导致上述衍射级光强强弱关系： $\{\tilde{E}_{3;m_x,m_y,m_z;l_x,l_y,l_z}\}$ 与 \tilde{E}_1, \tilde{E}_2 的能流密度的交叠程度，会因 $\{k_{3;m_x,m_y,m_z;l_x,l_y,l_z}\}$ 相对于 $k_1 + k_2$ 的偏离程度的增加而降低，导致在晶体内有效倍频距离下降，进而导致倍频效率降低，以至最终相应级次 m_x, m_y, m_z 的倍频光集的强度下降，导致远场成像平均光强降低。而 $\{k_{3;2,1,0;l_x,l_y,l_z}\}$ 比 $\{k_{3;1,1,0;l_x,l_y,l_z}\}$ 相对于 $k_1 + k_2$ 偏离得更严重，因此 $\{k_{3;2,1,0;l_x,l_y,l_z}\}$ 所对应的 $\{\tilde{E}_{3;2,1,0;l_x,l_y,l_z}\}$ 整体光强较低，其远场成像图案也更暗，说明利用高阶倒格矢进行准相位匹配和实现非线性全息成像也有一定的走离效应。综合二者便可解释 2 级以上的衍射光强随级次的增加很快降低。

4 总结与展望

4.1 总结

本文先调研了基于二维准相位匹配结构的二维非线性菲涅尔全息，接着围绕基于三维准相位匹配结构的三维非线性傅里叶全息，探讨了以下内容：

先从非线性电磁介质中麦氏方程组的微分形式导出电场波动方程的最普遍形式，接着通过单色化参与混频的复色电场及介质中电磁相关的复色线性和非线性极化场、分离时间项、引入电磁非线性效应，得到三维空间中的含磁非线性电介质中的定态单色电场波动方程的最普遍形式；同时通过附录 C 详细证明了当完美相位匹配条件满足时，在磁各向同性介质中，不论介质是否是电各向异性的，波动方程的解均随着波动方程右侧非线性波源项值的减小而收敛于单色定振幅平面电磁波，为二阶非线性系数与一阶非线性系数均不为零的磁各向同性铌酸锂单轴晶体内定态波动方程的试探解的相位部分保持平面波解、振幅采用缓变振幅近似的常数变易法提供了理论依据。

同时通过附录 C 的另一结论即电各向异性与电各向同性的磁各向同性非线性电介质中波动方程的解的形式相同，允许我们跳过各向异性的坎，直接在电磁各向同性的非线性电介质中的三维单色定态波动方程的基础上做文章；而为考虑波动方程的二阶和频波和倍频波解，方程右侧非线性波源项只保留二阶非线性极化过程生成的和频极化光，并利用克莱曼对称性、构建固定且特殊的实验几何场景和相应姿势的坐标系以计算有效二阶非线性系数、对波动方程两侧各矢量场保留相位地取模，得到标量场近似下的一维和三维的波动方程。接着由波动方程的标量形式导出描述三波混频的三个一维和三维的耦合波方程，给出了一维、三维完美相位匹配条件，以及相位不匹配时，随波矢失配的增加和频波振幅的振荡情况。

然后给出了非线性电介质的正常色散是波矢失配和完美相位匹配无法实现的原因，而为最大程度地利用铌酸锂晶体的最大二阶非线性系数张量分量 d_{33} 以及为实现倍频的同时还要成像，则无法使用铌酸锂晶体的双折射现象以弥补波矢失配量，为此必须采用准相位匹配技术。透过周期性电极化反转的一维准相位匹配结构的研究，发现结构的特定倒格矢可以补偿相应的波矢失配，以至于牺牲一点转换效率可以完成相位匹配过程；而考察飞秒激光直写的三维光子晶体中的准相位匹配机制，我们发现

$\chi_{eff}^{(2)}(\mathbf{r})$ 经三维周期性调制的光子晶体可以从三个维度上弥补和频波生成过程中的波矢失配，并且三维标量波动方程并未限制和频波波矢方向，只限制了其大小，导致所有终点落在和频波波矢旋转椭球面上的结构倒格矢均可弥补相应方向上的波矢失配，以至三维光子晶体可以同时簇射波矢方向不同的携带结构角谱信息的二阶和频光集，说明三维光子晶体具有实现三维非线性傅里叶全息的能力。

继证明了三维光子晶体可以实现三维非线性全息后，由于飞秒激光光斑大小暂无法自由调整和均一分布、和频光集携带的是结构角谱信息，以及远场成像相对于近场成像的优势，我们对三维非线性全息采取了傅里叶体全息技术，基于目标图像的远场光强分布，考虑给定的衍射级次所对应的倒格矢和细节倒格矢平面，匹配了基频光波长并初始化结构的三维角谱分布后，通过三维 GS 算法，逆向设计出了三维光子晶体的结构正空间相位分布。

为将所得的三维相位分布刻写为三维光子晶体的实际 $\chi_{eff}^{(2)}(\mathbf{r})$ 分布结构，我们利用三维迂回相位编码算法，将三维 GS 算法生成的纯相位体全息图，转换为 240 层中每层飞秒激光直写的光斑坐标 x 向偏移量阵列，并且由于坐标偏移量与相位的分级数为 30 级，实现了对三维光子晶体的多值化准连续编码。在这样的高坐标移动精度和正确的设计思路下，所编码完成的结构在远场数值模拟仿真中成功实现了准相位匹配过程即基频光到倍频光集的转换，并得到了倍频光集在三维光子晶体后厘米级光屏上的远场衍射所成目标像，即只利用同一个三维光子晶体，同时实现了二次谐波的生成和三维非线性全息编码成像，完成了三维非线性傅里叶计算体全息全过程。

4.2 展望

二阶非线性系数经空间调制的非线性光子晶体被广泛用于完成准相位匹配过程以实现激光变频，周期性极化的铌酸锂晶体便是其中最受欢迎的非线性光子晶体之一。自从 1998 年，非线性光子晶体的概念在理论上从一维被拓展到二维，许多有趣的非线性光学效应，如级联非线性光学过程、非线性切连科夫辐射、非线性 Talbot 自成像，以及量子路径纠缠等，在过去的二十年内被广泛地研究和报道了^[31-38]。

随着准相位匹配理论和飞秒激光直写等加工技术的发展，非线性光子晶体从二维拓展到三维，2018 年，最后一个空间维度的突破允许我们从三个维度上实现全方位的准相位匹配。基于二阶非线性系数在三个维度上均被调制的三维光子晶体，本文在完成三维准相位匹配过程的同时，实现了三维非线性全息编码成像，展示了三维光子

晶体在三维非线性远场傅里叶计算体全息成像方向上的应用之一。

本文在技术上可优化的地方至少还有两处，一是加工精度中的位移精度远远高于光斑大小的精度，即机械结构的最小位移步长远小于飞秒激光光斑最小直径。在本文的第三章中，三维迂回相位编码的逐层开孔的小孔面积均采用光斑最小面积，是为了实现加工平台所允许的最大像素尺寸 16×16 像素的目标物光场输出成像，为此只能用三维 GS 算法反向生成与目标图像的复振幅全息图效果相似的纯相位全息图，但毕竟飞秒激光光斑最小面积的不可变所导致的傅里叶全息图振幅部分的丢失，不可避免地会降低最终的远场重建像的成像质量。

为此，为了使得光斑面积可变，在光斑最小面积不变的情况下，可尝试在相同尺寸的总调制区域内设计更大的周期结构以开面积大于光斑面积的孔，使得大光斑面积为小光斑面积的交集，以生成可变面积的孔，这样便可不经三维 GS 算法生成纯相位全息图而直接采用目标物像的傅里叶变换设计三维光子晶体结构，以提高成像质量。

但这样会牺牲目标图像的像素即尺寸大小。因此接下来技术上应努力的方向是减小飞秒激光光斑直径的同时保持其能量密度不降低太多，使得相同的周期结构内擦除二阶非线性系数的程度即调制深度不变，但加工区域面积可变，以通过擦除区域中心位移量阵列调制正空间三维光子晶体内非线性全息图相位分布的同时，连同其三维振幅分布也一并调制，实现在光子晶体内直写目标物的三维复振幅体全息图。即使保持现有设计思路不变，减小飞秒激光光斑面积也可提高目标物像的像素尺寸和分辨率。

第二个可优化的目标点是在三维 GS 算法上，在叠次傅里叶变换的每次傅里叶变换到角谱分布的迭代之前对正空间中的复振幅全息图所加入的振幅分布限制不一定要全空间的分布都置 1，可以随着迭代次数的增加，循序渐进地从弱振幅限制到强振幅限制，这样所收敛到的纯相位全息图应更接近全局最优解，则其作用于重现倍频光后所还原得到的倍频光的正空间远场分布的成像质量理应会更好。

而在第三章的波长匹配一节，也暗示可以将不同目标图像分别置于不同级次的倒格矢所指向的倒格子，以及不同的细节倒格矢 $x-o-z$ 平面上，则可利用一次电光效应，对三维非线性晶体外加不同大小的静电场，改变其一阶非线性系数即折射率的大小，从而改变晶体内基频光的波长和波矢曲面半径，以至于匹配到不同级次的倒格矢和细节倒格矢，出射不同波长的同频同颜色倍频光集，在远场复原不同的目标重建像。

另外，在同时实现了三维准相位匹配和三维非线性全息成像之余，还可以考虑优化目标结构，在实现准相位匹配完成基频光到倍频光的转换的同时，完成对倍频光束

的整形，并用整形后的具有特定波前或波形的倍频光束，实现三维非线性全息成像。这样便更加体现出三维光子晶体全方位调控光场的潜力之大，不仅能分别实现多种不同类型的光场操纵功能，最难能可贵的是在同一块三维光子晶体内同时实现三种及以上的光场调控技术。当一晶多用达到一定程度后，经三维非线性光子晶体调制后的出射高度有序的非线性光场所蕴含的多维信息量将远远大于相同能量下，只蕴含振幅和相位信息的普通激光光场，预示着三维光子晶体在应用层面的广阔前景；而本文所研究的三维非线性全息只是三维非线性光子晶体在非线性全息方向上的应用之一。

当非线性光子晶体的调制维度和准相位匹配维度提高到三维后，以前受技术限制不能实施甚至无法想象的点子现在都可基于三维非线性光子晶体进行深深浅浅的尝试，这或多或少印证了安德森前辈的那句“More is different”，以及法拉第先生的“新生的婴儿”一说。

参考文献

- [1] Maiman T H . Stimulated optical radiation in ruby[J]. *Nature*, 493–494 (1960).
- [2] Franken P A, Ward J F Optical harmonics and nonlinear phenomena[J]. *Reviews of Modern Physics*, 1963, 35(1): 23.
- [3] Armstrong, J. A., Bloembergen, N., Ducuing, J. & Pershan, P. S. Interactions between light waves in a nonlinear dielectric. *Phys. Rev.* **127**, 1918–1939 (1962).
- [4] Miller R C. Optical second harmonic generation in piezoelectric crystals[J] *Applied Physics Letters*, 1964, 5(1): 17-19.
- [5] Myers L E, Bosenberg W R, Eckardt R C, et al. Multimigrating quasi-phase-matched optical parametric oscillator in periodically poled LiNbO₃[J]. *Optics letters*, 1996, 21(8): 591-593.
- [6] Myers, L. E. et al. Quasi-phase-matched optical parametric oscillators in bulk periodically poled LiNbO₃. *J. Opt. Soc. Am. B* **12**, 2102–2116 (1995).
- [7] Zhu, S., Zhu, Y. Y. & Ming, N. B. Quasi-phase-matched third-harmonic generation in a quasi-periodic optical superlattice. *Science* **278**, 843–846 (1997).
- [8] Chen, B. Q., Zhang, C., Hu, C. Y., Liu, R. J. & Li, Z. Y. High-efficiency broadband high-harmonic generation from a single quasi-phase-matching nonlinear crystal. *Phys. Rev. Lett.* **115**, 083902 (2015).
- [9] Xu, T. X. et al. Three-dimensional nonlinear photonic crystal in ferroelectric barium calcium titanate. *Nat. Photonics* **12**, 591–595 (2018).
- [10] Xu, T. et al. A naturally grown three-dimensional nonlinear photonic crystal. *Appl. Phys. Lett.* **108**, 051907 (2016).
- [11] Berger, V. Nonlinear photonic crystals. *Phys. Rev. Lett.* **81**, 4136-4139 (1998).
- [12] Kroesen, S., Tekce, K., Imbrock, J. & Denz, C. Monolithic fabrication of quasi phase-matched waveguides by femtosecond laser structuring the $\chi^{(2)}$ nonlinearity. *Appl. Phys. Lett.* **107**, 101109 (2015).

- [13] Wei, D., Wang, C., Wang, H. et al. Experimental demonstration of a three-dimensional lithium niobate nonlinear photonic crystal. *Nature Photon* **12**, 596–600 (2018).
- [14] Wei, D., Wang, C., Xu, X. et al. Efficient nonlinear beam shaping in three-dimensional lithium niobate nonlinear photonic crystals. *Nat Commun* **10**, 4193 (2019).
- [15] Keren-Zur, S., Ellenbogen, T. A new dimension for nonlinear photonic crystals. *Nature Photon* **12**, 575–577 (2018).
- [16] X. Hong, B. Yang, C. Zhang, Y Qin, and Y. Zhu. "Nonlinear Volume Holography for Wave-Front Engineering," *Phys. Rev. Lett.* **113**. 163902 (2014).
- [17] Zhang, J., Zhao, X., Zheng, Y., Li, H. & Chen, X. Universal modeling of second-order nonlinear frequency conversion in three-dimensional nonlinear photonic crystals. *Opt. Express* **26**, 15675–15682 (2018).
- [18] Bahabad, A. & Arie, A. Generation of optical vortex beams by nonlinear wave mixing. *Opt. Express* **15**, 17619–17624 (2007).
- [19] Lee, W.-H. Binary computer-generated holograms. *Appl. Opt.* **18**, 3661–3669 (1979).
- [20] Kroesen, S., Tekce, K., Imbrock, J. & Denz, C. Monolithic fabrication of quasi phase-matched waveguides by femtosecond laser structuring the $\chi^{(2)}$ nonlinearity. *Appl. Phys. Lett.* **107**, 101109 (2015).
- [21] S. Janz *et al.*, *Appl. Phys. Lett.* **65**, 216 (1994).
- [22] J. P. Bouchard *et al.*, *Appl. Phys. Lett.* **77**, 4247 (2000).
- [23] K. Zeaiter *et al.*, *Opt Lett.* **28**, 911 (2003).
- [24] D. Artigas, E. U. Rafailov, P. Loza-Alvarez & W. Sibbett, *IEEE J. Quantum Electron.* textbf{40}, 1122 (2004).
- [25] P. G. Kazansky, V. Pruneri & P. S. J. Russell, *Opt. Lett.* **20**, 843 (1995).
- [26] A. Rose & D. R. Smith, *Phys. Rev. A* **84**, 013823 (2011).
- [27] Magel, G. A., Fejer, M. M. & Byer, R. L. Quasi-phase-matched second-harmonic generation of blue light in periodically poled LiNbO₃. *Appl. Phys. Lett.* **56**, 108–110 (1990).
- [28] Broderick, N. G., Ross, G. W., Offerhaus, H. L., Richardson, D. J. & Hanna, D. C.

- Hexagonally poled lithium niobate: a two-dimensional nonlinear photonic crystal.
Phys. Rev. Lett. **84**, 4345–4348 (2000).
- [29] Dan Wei, Jiale Guo, Xinyuan Fang, Dunzhao Wei, Rui Ni, Peng Chen, Xiaopeng Hu, Yong Zhang, Wei Hu, Y. Q. Lu, S. N. Zhu, and Min Xiao, "Multiple generations of high-order orbital angular momentum modes through cascaded third-harmonic generation in a 2D nonlinear photonic crystal," *Opt. Express* **25**, 11556-11563 (2017).
- [30] Arie, A., Habshoosh, N. & Bahabad, A. Quasi phase matching in two-dimensional nonlinear photonic crystals. *Opt Quant Electron* **39**, 361–375 (2007).
- [31] Broderick, N. G., Ross, G. W., Offerhaus, H. L., Richardson, D. J. & Hanna, D. C. Hexagonally poled lithium niobate: a two-dimensional nonlinear photonic crystal.
Phys. Rev. Lett. **84**, 4345–4348 (2000).
- [32] Saltiel, S. M. et al. Generation of second-harmonic conical waves via nonlinear bragg diffraction. *Phys. Rev. Lett.* **100**, 103902 (2008).
- [33] Zhang, Y., Wen, J., Zhu, S. N. & Xiao, M. Nonlinear Talbot effect. *Phys. Rev. Lett.* **104**, 183901 (2010).
- [34] Jin, H. et al. Compact engineering of path-entangled sources from a monolithic quadratic nonlinear photonic crystal. *Phys. Rev. Lett.* **111**, 023603 (2013).
- [35] Lu, R. E. et al. Nearly diffraction-free nonlinear imaging of irregularly distributed ferroelectric domains. *Phys. Rev. Lett.* **120**, 067601 (2018).
- [36] Fang, X. et al. Coupled orbital angular momentum conversions in a quasi-periodically poled LiTaO₃ crystal. *Opt. Lett.* **41**, 1169–1172 (2016).
- [37] Zhang, Y., Gao, Z. D., Qi, Z., Zhu, S. N. & Ming, N. B. Nonlinear Cerenkov radiation in nonlinear photonic crystal waveguides. *Phys. Rev. Lett.* **100**, 163904 (2008).
- [38] Chen, C. et al. Cerenkov third-harmonic generation via cascaded $\chi^{(2)}$ processes in a periodic-poled LiTaO₃ waveguide. *Opt. Lett.* **36**, 1227–1229 (2011).

致 谢

衷心感谢张勇老师和高志达教授在我毕设过程中分阶段提供了各种可行的方向性指导和细节性建议，促进我兼顾效率与质量地完成了毕设任务。

在这里必须要感谢和祝福有幸言传身教于的东北大学理学院物理系任课老师们，如果这个系统没有你们的辛勤耕耘，作为相同的输入的我，便不会有今天的输出。我的三维血肉在每一个方向的投影都蕴含了这些教师的身影，他们包括：李楠、朱林、柳艳、胡勇、王旗、张建峰、公卫江、吴海娜、刘国良、高志达、杨玉玲、孟祥颖、李云鹤、何砚发、代雪峰、张德安、易光宇、白质明、张敬飞，也希望大家保持教学热情，为应用物理学、基础物理学培养出更多壮苗。

在整个本科生涯，东北大学理学院其他系以及其他学院的教师也曾极大地震撼过我的神经末梢，他们是：理学院数学系教师郭阳、孔庆海、张雪峰，化学系教师孙宏滨、王卓鹏，软件学院教师高天寒，信息科学与工程学院教师李景宏、马大中、孙玉舸。同时也感谢陶天琦、岳登晓两位辅导员老师在四年里对我们的成长呵护有加。

感谢三年同寝舍友翟心哲、周健豪、姚家铭前仆后继地左右为男、上下其手，锻炼和培养了我中通外直的内禀属性，同时也带给了我苦尽甘来的欢声笑语。

在本文的撰写过程中，陈鹏程师兄无私地给予了我很大的程度的理论指导和技术支持。这是我第一次系统地主动接受外界的影响和参与较强的相互作用，收获颇丰。

杨振宁在他的《六十八年心路》中提到，当下物理学所面临的困难有增无减的原因，在于随着实验物理与理论物理的成熟和发展，这两个领域的物理实验均演变得越来越复杂、成本和周期都越来越不堪重负，一方面导致理论物理学家与理论物理学之间的鸿沟，甚至跨越了实验物理学家与理论物理学家之间的鸿沟，另一方面导致理论物理研究生，甚至连同实验物理研究生，都离实验现象越来越远；但对物理学科本身，问渠那得清如许，就是因为它是一门基于实验的学科。但学科在发展，有序度在集成，复杂度在提升，就像不可逆的熵增，物理学的实验现象离接下来一代代学生渐行渐远已经大势所趋，再加上上个世纪费曼在巴西的考察所经历的噩梦一遍又一遍地正在中国和其他国家发生，以至于梦寐以求的大统一理论可能在 22 世纪之前仍无法掌握。若朝闻道的曙光在我有生之年确定无法一瞥，那我宁可不去飞蛾扑火地朝思暮想它。

以上这段文字为在接下来的研究生生涯中，希望近距离接触物理特别是能动手实验的我，在保研时选择专业方向提供了指导思想，那就是该专业首先必须要存在实验，不能是纯理论，否则我无法通过现象来否定或肯定自身的理论，以及评判、分析或查看他人的理论的过人之处和弊病之处，也就会陷入大学时期的天马行空的纯理论泥潭，尽管沉浸于自己创造的自洽逻辑中也很有趣，但对物理而言终究是不完整的；其次，实验成本必须相对低廉（尽管高端实验都贵且偏向自动化）、必须触手可及、必须有人工操作的部分、实验平均周期也不能过长。这样光学方向就至少满足条件了。

另外，在杨的盛宴已过和安德森的多者异也的“抛砖引玉”下，我发现在极小和极大尺度的领域，物理学中美丽的女人都已经被爱因斯坦、朗道那个世纪的年轻人所占有了，我既不太愿意或许也没有能力和勇气甚至没有可能性，在没有照顾好自己的同时去处理这些天才们剩下的旁枝末节；所以我费尽心思说服自己必须选择中间尺度的舞台，一方面补上我所欠缺的实验技巧，另一方面也给自己一个机会去重新认识和还原人眼所能接受的尺度的量级下，自然的本来面目；而处于中间尺度的物理学圣地凝聚态物理，又过于人才辈出、诸神之战，独木不成林的我必须选择一个颇有建树的人作为导师，但矛盾在于鉴于本科的经验，我提出的理论很可能与大多数人迥异，若与导师无法兼容便可能成为灭顶之灾，特别是当导师的理论被反复验证有考虑不周之处时。所以我找不出理由说服自己迈出选择极小或极大，以及凝聚态物理的那一步。

在这些矛盾下，我暂时远离了我从课本中熟悉到的每一个旧领域，选择了另一个比原子尺度即纳米级的凝聚态物理体系更大一点的，光波长级即微米级的另一个更“中间”一点的中间尺度体系——非线性光子晶体领域。毕竟我需要一个肉眼可见的顷刻间就能决定理论是对是错的现象级领域，在那里任何人所提出的每一个理论首先必须是可检验的，即既可证实也可证伪，其次检验的时间成本也得相对较低，在这样的思想下我选择加入张勇老师在 2018 年刚开创的新领域中，暂时免去了与物理学中几大领域直接竞争的局面。这里虽然离现象更接近而在理论物理学家看来不够基本，但只要够“正确”即可，即只要每一步都在做正功，总会越来越接近还原事情的真相。

既然极小尺度与极大尺度充满了智者和历史的尘埃且至今仍然乌云不散，可能上帝正是在以这种碰壁的方式启发我们回过头来从与自身尺度相符的领域找寻那把丢失的钥匙；不过即使这个问题在该尺度或根本无解也没关系，因为在这个领域恰好可以不用去担心上帝允不允许相应事情的发生，也不用关心上帝会以什么样的奇淫巧技

实现这样的过程，反而可以在不违背他老人家的基本意愿的前提下，为作品添加个人的自由意志，兼顾科学和艺术地成为该新领域的国王，而不必把重心放在那位创生我们的普罗米修斯身上。毕竟当后代都开始按照我们的新规则运作或熟悉我们的新理论时，便是我们的理论超越其后辈和同辈以及它对它自身的约束之时——在年轻的新领域独树一帜的概率和难易程度，等价于在旧宇宙里创造新宇宙本身，相对于思考如何修缮、维系甚至认识庞大的旧宇宙的难易程度。所以我选择了将精力暂时用于培养和创造、规范新的下一代，而不是冒着风险找出上一代的瑕疵并说服和修正它。

尽管光学非线性效应从其公式确实看不出除了能量守恒外更深层次的物理意义，但 Landau、Butcher、Lifshitz、Pershan 等人确实已经对各阶非线性系数完成了数学描述并为其在量子理论中找到了安家之地，有了这点我就不必过于担心了；相反，非线性过程从表达式来看一般人也都望而却步，这对于将要与那些年轻人同台竞技的我而言倒是优点。我既不畏惧艰深但优美的理论构建，又不嫌弃眼花繁杂的数值演算，也不排斥动手实验，毕竟所有道路都通向罗马，都叫格物致知，而这三条道路也是去往罗马的必经之路，少了一条都像华容道一样虽然仍能封闭和自洽但局外人看还是缺了块拼图。我爷爷说，青年在选择领域时的考虑就像娶老婆一样，回头看才会发现，只要她心坎好、合得来，哪哪儿和谁谁儿都一样；再加上旧有领域虽美但精致而森严的妆容让人不寒而栗，这样我就直接选了一个此前从未听闻过的年轻的新领域。

每个浪漫的男科学家都喜欢独坐幽篁之时，大道至简、指点江山地竹中窥落日，但历史腾出的这样的机会已经不多了，我这样的个体在时下的大环境下大约只可以在粗茶淡饭之余借闲情逸致将其推演一番，尽管在理论物理学家看来我们的领域里的排列组合才更像把玩之物……但每一个认真严格且快乐地对待自己的领域和作品并与之一同前行的人，又何尝不是在各自的时空长廊里研究和摆设自己的新旧玩具呢？

最后，引用《球状闪电》里的两段文摘，“理想主义者和玩世不恭的人都觉得对方很可怜，可他们实际都很幸运。”幸运地能够活在了自己的世界。而我们踌躇在应用和理论的边缘，一面是理论的桃花园林，一面是技术的屋舍俨然，既然一只脚触摸到了现实，就不能假装没有看到这份现实。——“金黄色的树林里分出两条路，可惜我们不能同时去涉足，但我们却选择了，人迹罕至的那一条，这从此决定了我们的一生。”林云喃喃地吟起了弗罗斯特的那首诗，声音像林间的清泉。

I'll go gentle into that good night.

附录 A 中文译文

一种基于双向误差扩散的数字菲涅尔全息图向纯相位全息图转换的新方法

P.W.M. Tsang^{1,*} and T. -C. Poon^{2,3}

摘要：我们报告了一种新颖、快速的数字转换方法，可将复振幅的菲涅耳全息图转换为纯相位全息图。简单来说，逐行地依次扫描复数全息图中的像素，且从相反的方向扫描奇数行和偶数行，形成双向误差扩散过程。每次访问的幅度像素被强制为恒定值，同时保留精确的相位值。产生的误差扩散到未曾被访问过的像素点中去。由此产生的新颖的纯相位全息图称为双向误差扩散（BERD）全息图。来自 BERD 全息图的重建图像，保真度高于那些从原始的复振幅全息图获得的重现像。

1. 引言

过去二十年里，计算机全息术（CGH）经历了令人鼓舞的发展。导致 CGH 成功的主要因素之一，可归功于极大加快了数字全息图的计算速度的快速算法的出现。最重要的是，计算和半导体技术的飞速发展，还使得廉价的商用 PC 和硬件（比如可编程逻辑门阵列 FPGA），能够以数字方式很快地生成中等尺寸的数字全息图。

尽管结果令人满意，但全息图显示还是一个困难的问题，这在一定程度上对全息技术的实际实现施加了瓶颈。主要问题之一是，像以硅为衬底的液晶（LCoS）那样能够显示全息图的高分辨率设备，却只能要么显示复振幅全息图的幅度，要么显示其相位分量，没法二者同时都显示。

对这个问题的一种直接解决方案是，采用两个空间光调制器（SLMs），来显示实部和虚部两分量，或者说复振幅全息图的振幅和相位信息。同样，复振幅全息图可以用一对纯相位全息图模拟，并用一对纯相位 SLMs 显示。

在一些方案中，将一对 SLMs 替换为单个设备，显示一对全息图；然后在之后，通过光栅合并重建的波前。尽管这种方法是有效的，但是光学装置的安装和排列相当复振幅。此外，分配给全息图的两个分量中每个分量的，SLM 的面积，会减少到其

原始面积的一半。

另外，也可以将复振幅全息图转换为纯振幅或纯相位全息图（POH），以便可以用单个设备直接显示。但是，如果使用纯相位显示方式，来显示菲涅耳全息图，则重建的图像将被散焦的两个孪生像污染。通过从同轴全息转换为离轴全息，孪生像中的那个实像，可以与重现的虚像分离开，然而两个图像之间的角度分离程度，受到显示器的分辨率的限制。

此外，由于干涉条纹图案的不透明性，减弱了重现光的透射光强，因此振幅全息图的光学效率相当低。另一方面，一个 POH 可以用纯相位设备显示，从而获得更高的光学效果效率，并同时消除孪生像。不利的一面是，消除(实)振幅分量，将导致重建图像严重失真(变形)。

为了减轻这个问题，通常采用 Gerberg-Saxton 算法，或叠次菲涅耳变换，这两个迭代的方法，来计算相位全息图，从而重建图像将与目标平面图像匹配。但是，用这种方法生成的相位全息图，是需要消耗大量计算资源的。

迄今为止开发的生成 POH 的最快方法，很可能是“一步式相位恢复（OSPR）过程”。简而言之，首先在生成数字全息图之前，就将随机相位(片)添加到物平面的每个(像素)点上。全息图的相位分量，可以被阈值量化或误差扩散后，用纯相位设备显示。

由 OSPR 过程生成的重建图像通常有噪声，并且有多个子帧，每个子帧代表 添加了不同的随机相位模式 的相同的物平面，并必须快速呈现给观察者，以平均掉/抹平散斑噪声。然而，子帧需要更多的计算来生成，而且还需要以高帧速率显示，以避免闪烁。

在本文中，我们提出了一种新颖而快速的方法，将复振幅菲涅耳全息图转换成单一相位的全息图，称为双向误差扩散全息图。生成的 BERD 全息图能够表示物平面，并在重建图像上保持良好的视觉质量。与 OSPR 方法或其他从物平面生成纯相位全息图的迭代方法不同，我们提出的方法可以直接应用于给定的数字复振幅全息图。

2. 一种新奇的纯相位全息图：双向误差扩散全息图

2.1 基于单向误差扩散的基本方法

首先，考虑一个复振幅的菲涅耳全息图，它是由物平面($x-o-y$ 面)上每个点发出的物光波生成的。全息图中的每个像素点用 $P(u,v)$ 表示，并且有相应的复振幅分布 $H(u,v)$ (其中 u,v 分别是全息图像面的坐标系的垂直轴和水平轴，见 Fig. 1(a))，该未经处理的菲涅尔全息图的表达式由下式给出：

$$H(u,v) \Big|_{\substack{0 \leq u < X \\ 0 \leq v < Y}} = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \frac{I(x,y) \exp(i2\pi r_{x,y;u,v} / \lambda)}{r_{x,y;u,v}}, \quad (1)$$

其中的 $I(x,y)$ 与 $r_{x,y;u,v} = [(u-x)^2 \delta + (v-y)^2 \delta + w_{x,y}^2]^{0.5}$ 分别是物平面上 (x,y) 点的光强，以及它距离全息图像面上 (u,v) 点的距离。物平面上某点 (x,y) 距离全息面的距离，用 $w_{x,y}$ 表示； λ 是光波的波长， δ 是全息面的单个像素的尺寸(边长)。 X 和 Y 是全息图(像素矩阵)的行数和列数，并且假定与物平面的尺寸一致(包括单个像素点的尺寸，以及像素点的行数和列数)。

我们假设在全息图的产生过程中没有加入随机相位。复振幅全息图可以转换成一个纯相位全息图 POH，即 $H_p(u,v)$ ，通过将每个像素点的 $H_p(u,v)$ 的(实)振幅值 $|H_p(u,v)|$ 设为 1，同时保持 $H_p(u,v)$ 的相位与 $H(u,v)$ 的相位一致：

$$|H_p(u,v)| = 1, \text{ and } \arg(H_p(u,v)) = \arg(H(u,v)). \quad (2)$$

然而，我们将在后面指出，使用 Eq. (2)生成的 POH，其重现图像的质量非常差。显然，重建图像的严重失真是由于去除振幅信息后，全息图中每个像素点的误差都很大。

针对上述问题，我们提出使用 Floyd-Steinberg 误差扩散技术，来补偿每个全息图像素的误差。该过程概述如下。首先，将全息图中的每个像素点按从左到右的顺序逐行扫描，将待求值的像素点的值，根据式(2)转换为纯相位量 $H_p(u,v)$ 。假设 $P(u_j,v_j)$ 是当前正在处理的像素。通过将振幅简单设定为单位 1，将会产生一个误差：

$$E(u_j,v_j) = H(u_j,v_j) - H_p(u_j,v_j). \quad (3)$$

接着，由单纯地将振幅置一以只取相位所产生的该误差 $E(u_j,v_j)$ 将被扩散到之前没有访问过的邻域像素。它的邻域成员的复振幅的值 H 将根据以下公式进行更新：

$$H(u_j, v_j + 1) \leftarrow H(u_j, v_j + 1) + w_1 E(u_j, v_j), \quad (4)$$

$$H(u_j + 1, v_j - 1) \leftarrow H(u_j + 1, v_j - 1) + w_2 E(u_j, v_j), \quad (5)$$

$$H(u_j + 1, v_j) \leftarrow H(u_j + 1, v_j) + w_3 E(u_j, v_j), \quad (6)$$

$$H(u_j + 1, v_j + 1) \leftarrow H(u_j + 1, v_j + 1) + w_4 E(u_j, v_j), \quad (7)$$

其中运算符"←"表示用表达式右侧的变量更新(覆盖)表达式左侧的变量(的值)。常系数 $w_1 \sim w_4$ 的值分别被设置为文献[16]中所采用的值: 7/16、3/16、5/16、1/16。

根据方程式(4)~(7), 如图 1(b)所示为 $P(u, v)$ 和它的邻域像素通过扩散来补偿误差的空间关系图中显示了 $P(u_j, v_j)$ 是怎样将误差扩散到它的四个未访问的邻域像素的。等价地, 方程式(4-7)可以如下地重新排列成一个紧凑的递归表达式:

$$H(u_j, v_j) \leftarrow H(u_j, v_j) + w_1 E(u_j, v_j - 1) + w_4 E(u_j - 1, v_j - 1) + w_3 E(u_j - 1, v_j) + w_2 E(u_j - 1, v_j + 1), \quad (8)$$

根据式(8), 图 1(c)形象地显示了每个像素更新后的值可以看做为其四个相邻像素的加权和。当全息图的所有像素被访问和处理后, 原始的全息图将被转换成一个纯相位全息图 POH。我们将在仿真实验部分说明, 用该方法得到的全息图 $H_p(u, v)$ 的重现像, 与用原始复振幅全息图得到的重现像相似。

2.2 基于双向误差扩散过程的增强方法

2.1 节所述的基本方法虽然有效, 但会对重建图像造成噪声污染。一种可能的解释是, 在误差扩散过程中, 如式(8)所示, 每个全息图像素所掺入的都是与之相对位置固定的各相邻像素所扩散的误差之和。这构成了一个类似于可预测的编码器在编码误差的过程, 在这个过程中产生的误差在一定程度上与信号相关。因此, 全息 $H_p(u, v)$ 的重构图像, 将与相关的与误差有关的噪声信号叠加在一起形成最终的重建像。

为了克服这一问题, 我们提出了一种增强的方法来降低全息信号与误差之间的相关性。首先, 我们注意到在 4~7 号方程中表示的误差扩散过程与扫描全息图图像中每一行像素的方向无关。因此, 我们改变了每一行像素的扫描方向, 使得 Eq. (8)中的递归过程, 在每一行像素被访问后都会中断, 从而导致了相关的误差, 在每一行被访问完毕后会被打断 (而不是一直积累)。

奇数行产生的误差以 Eqs. (4-7)中给出的方法扩散到相邻的像素，而偶数行则通过反向(从右到左)应用误差扩散过程进行补偿，如下面的 Eqs. (9-12)所示。像素 $P(u_j, v_j)$ 及其邻域像素之间的空间关系如图 1(d)所示。

$$H(u_j, v_j - 1) \leftarrow H(u_j, v_j - 1) + w_1 E(u_j, v_j), \quad (9)$$

$$H(u_j + 1, v_j + 1) \leftarrow H(u_j + 1, v_j + 1) + w_2 E(u_j, v_j), \quad (10)$$

$$H(u_j + 1, v_j) \leftarrow H(u_j + 1, v_j) + w_3 E(u_j, v_j), \quad (11)$$

$$H(u_j + 1, v_j - 1) \leftarrow H(u_j + 1, v_j - 1) + w_4 E(u_j, v_j). \quad (12)$$

由基于处理复振幅全息图的偶数行的 Eqs. (4-7)，以及处理奇数行的 Eqs. (9-12)，所产生的全息图，称为双向误差扩散(BERD)全息图。

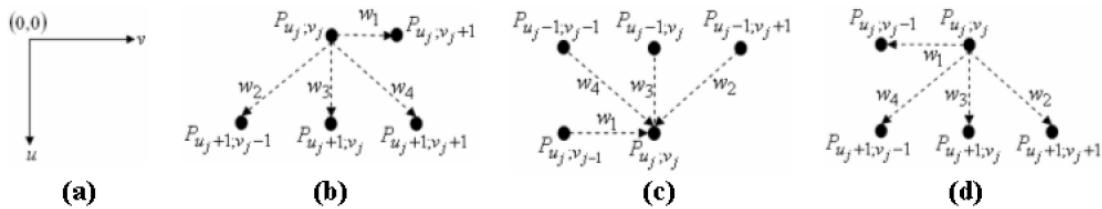


Fig. 1. (a) Horizontal and vertical axes of the co-ordinate system. (b) Diffusion of the error from the current pixel to its neighborhood, scanned from left to right. (c) An alternative representation of (a), showing the updating of a pixel from its neighborhood. (d) Diffusion of the error from the current pixel to its neighborhood, scanned from right to left.

3. 实验结果

采用二值化图像“CITYU”、平滑图像“Lena”，和高纹理图像“Mandrill”（如图 2(A)-2(c)所示）对本文方法进行评价。应用式(1)生成三幅图像的复振幅菲涅耳全息图。其中，全息图以及（作为物像）测试的图像，都包含 2048×2048 个像素，每一个像素都是 $7\mu\text{m} \times 7\mu\text{m}$ 的方形尺寸。

光束的波长为 650 nm 。所有的图像都平行于全息图平面，且位于距全息图平面 0.3m 处。这三种复振幅全息图的数值重建图像与原始图像相同，因此在这里没有显示。

接着利用式(2)简单去除全息图的实振幅分量，只保留菲涅尔全息图的相位部分，数值重建图像如图 3(a)-3(c)所示。可以看到，这三幅重建图像的阴影区域(的原图信息)被广泛地删除了，留下的大部分是图像的边缘和轮廓。

为了克服这一问题，我们在 2.1 节中提出的单向误差扩散方法被应用于将复振幅

全息图转换为纯相位全息图（POHs）。重建的纯相位全息图 POHs 所对应的图像如图 4(a)-4(c)所示。我们观察到重建的图像与图 2(a)-2(c)中的原始图像很接近。

然而，也注意到一定数量的噪音污染，尤其是在黑暗区域。

随后，我们用我们提出的双向误差扩散方法将复振幅全息图转换成 BERD 全息图，如 2.2 节所述。如图 5(a)-5(c)所示，用这种方法生成的 BERD 全息图的数值重构图像与原始图像非常相似，并且噪声信号不明显。

将由上述三种方法得到的 POHs 重建图像的质量，与用复振幅全息图得到的 POHs 重建图像的质量，通过保真度评价方法进行定量评价，如表 1 所示。显然，所提出的双向误差扩散方法所生成的 BERD 全息图的保真度(在 PSNR 方面)明显高于同类算法所得的全息图。为清晰起见，图 2-5 以它们的实际大小分别显示在媒体文件 MF2-MF5 中。



Fig. 2. (a)-(c) Test images “CITYU”, “Lena”, and “Mandrill”.

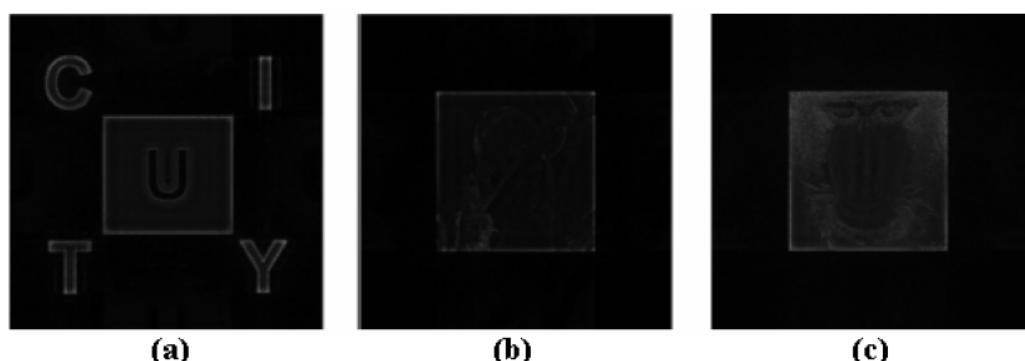


Fig. 3. (a)-(c) Reconstructed images of the phase-only holograms derived from direct removal of the magnitude component of the complex holograms.



Fig. 4. (a)-(c) Reconstructed images of the phase-only holograms derived from our proposed unidirectional error diffusion method.



Fig. 5. (a)-(c) Reconstructed images of the BERD holograms (novel phase-only holograms derived from our proposed bidirectional error diffusion method).

Table 1. Quantitative comparison between the fidelity of the reconstructed images as shown in the Figs. 3(a)-3(c), 4(a)-4(c), and 5(a)-5(c) (with reference to the ones obtained from the original complex holograms in Figs. 2(a)-2(c)) from phase-only holograms derived from different means. The reported values are PSNR values.

	CITYU	Lena	Mandrill
Direct removal of the magnitude component	13.0 dB	16.1 dB	16.2 dB
Unidirectional error diffusion	24.3 dB	28.1 dB	27.0 dB
Bidirectional error diffusion	31.6 dB	38.6 dB	32.7 dB

4. 结论

本文提出了一种快速将复振幅的菲涅尔数字全息图转换为纯相位全息图的新方法，即 BERD 全息图。这一过程是通过 BERD 机制完成的，每个全息图像素被依次扫描并强制约束为一个恒定的振幅，同时产生的误差扩散到未访问的邻域。奇数行和偶数行从相反的方向扫描，构成一个双向误差扩散过程。

我们提出的方法有三个主要优点。

首先，生成的新型 BERD 全息图的重构图像与用原始复振幅全息图得到的结果非常相似，证明纯相位全息图相对于复振幅全息图的近似是有用的。其次，我们所提

出的方法只涉及了中等的计算量，对于大部分图片一般像素和尺寸的图片，生成其所对应的纯相位菲涅尔全息图所消耗的计算资源并不是很大。第三，我们的方法可以直接应用于现有的复振幅菲涅耳全息图，而不需要在真实空间实际存在客观的物像。

一个潜在的研究方向是探索一套更好的误差扩散系数($w_1 \sim w_4$)，以及误差扩散矩阵元个数及其方向甚至非邻域扩散，以提高重建后的 BERD 全息图像的质量；还可以进行边界条件的优化，比如在扩散到最后一行之时，可将最后一行的误差扩散到第一行，同理，也可将最后一列的误差扩散到第一列，实现周期性边界条件下的单双向误差扩散法；也可以在左右双向误差扩散的基础上，加上上下方向后实行四向误差扩散算法，预期这样会进一步提高相应四向误差扩散算法所生成的纯相位全息图的重建像成像质量。

致谢

这项工作得到了中国科学院客座教授的支持，在高级国际科学家项目下的资助编号为 2010T2G17，和高端外国专家来华招聘项目，资助号 GDJ20130491009。

附录 B 外文原文

Novel method for converting digital Fresnel hologram to phase-only hologram based on bidirectional error diffusion

P.W.M. Tsang^{1,*} and T. -C. Poon^{2,3}

¹*Department of Electronic Engineering, City University of Hong Kong, Hong Kong China*

²*Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061, USA*

³*Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, P.O. Box 800-211, Shanghai 201800, China*

^{*}*eewmtsan@cityu.edu.hk*

Abstract: We report a novel and fast method for converting a digital, complex Fresnel hologram into a phase-only hologram. Briefly, the pixels in the complex hologram are scanned sequentially in a row by row manner. The odd and even rows are scanned from opposite directions, constituting to a bidirectional error diffusion process. The magnitude of each visited pixel is forced to be a constant value, while preserving the exact phase value. The resulting error is diffused to the neighboring pixels that have not been visited before. The resulting novel phase-only hologram is called the bidirectional error diffusion (BERD) hologram. The reconstructed image from the BERD hologram exhibits high fidelity as compared with those obtained with the original complex hologram.

©2013 Optical Society of America

OCIS codes: (090.0090) Holography; (090.1995) Digital holography; (090.1760) Computer holography.

References and links

1. J. Weng, T. Shimobaba, N. Okada, H. Nakayama, M. Oikawa, N. Masuda, and T. Ito, "Generation of real-time large computer generated hologram using wavefront recording method," *Opt. Express* **20**(4), 4018–4023 (2012).
2. P. Tsang, W. K. Cheung, T.-C. Poon, and C. Zhou, "Holographic video at 40 frames per second for 4-million object points," *Opt. Express* **19**(16), 15205–15211 (2011).
3. M. Makowski, A. Siemion, I. Ducin, K. Kakarenko, M. Sypek, A. Siemion, J. Suszek, D. Wojnowski, Z. Jaroszewicz, and A. Kolodziejczyk, "Complex light modulation for lensless image projection," *Chin. Opt. Lett.* **9**, 120008-1–120008-3 (2011).
4. M.-L. Hsieh, M.-L. Chen, and C.-J. Cheng, "Improvement of the complex modulated characteristic of cascaded liquid crystal spatial light modulators by using a novel amplitude compensated technique," *Opt. Eng.* **46**(7), 070501 (2007).
5. R. Tudela, E. Martín-Badosa, I. Labastida, S. Vallmitjana, I. Juvells, and A. Carnicer, "Full complex Fresnel holograms displayed on liquid crystal devices," *J. Opt.* **A5**, S189–S194 (2003).
6. C. K. Hsueh and A. A. Sawchuk, "Computer-generated double-phase holograms," *Appl. Opt.* **17**(24), 3874–3883 (1978).
7. X. Li, Y. Wang, J. Liu, J. Jia, Y. Pan, and J. Xie, "Color holographic display using a phase-only spatial light modulator," *OSA Technical Digest*, paper DTh2A.3 (2013).
8. S. Reichelt, R. Häussler, G. Fütterer, N. Leister, H. Kato, N. Usukura, and Y. Kanbayashi, "Full-range, complex spatial light modulator for real-time holography," *Opt. Lett.* **37**(11), 1955–1957 (2012).
9. H. Song, G. Sung, S. Choi, K. Won, H. S. Lee, and H. Kim, "Optimal synthesis of double-phase computer generated holograms using a phase-only spatial light modulator with grating filter," *Opt. Express* **20**(28), 29844–29853 (2012).
10. J.-P. Liu, W. Y. Hsieh, T.-C. Poon, and P. W. M. Tsang, "Complex Fresnel hologram display using a single SLM," *Appl. Opt.* **50**(34), H128–H135 (2011).
11. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik (Stuttg.)* **35**, 237–246 (1972).
12. J. Yeom, J. Hong, J. -H. Jung, K. Hong, J.-H. Park, and B. Lee, "Phase-only hologram generation based on integral imaging and its enhancement in depth resolution," *Chin. Opt. Lett.*, **9**(12), 12009-1–12009-4 (2011).
13. E. Buckley, "Holographic laser projection technology," *Proc. SID Symp.*, 1074–1078 (2008).

-
14. A. J. Cable, E. Buckley, P. Marsh, N. A. Lawrence, T. D. Wilkinson, and W. A. Crossland, "Real-time binary hologram generation for high-quality video projection applications," SID International Symposium Digest of Technical Papers, **35**, 1431–1433 (2004).
 15. E. Buckley, "Real-time error diffusion for signal-to-noise ratio improvement in a holographic projection system," *J. Disp. Tech.* **7**(2), 70–76 (2011).
 16. R. W. Floyd and L. Steinberg, "An adaptive algorithm for spatial grey scale," *Proc Soc. Info. Disp.* **17**, 75–77 (1976).
-

1. Introduction

Computer-generated holography (CGH) has undergone encouraging development in the past two decades. One of the major factors leading to the success in CGH can be attributed to the emergence of fast algorithms that speed up the computation of the digital hologram by a significant amount. On top of that, the rapid advancement of computing and semiconductor technologies, have also enabled medium size digital hologram to be generated numerically swiftly with low cost commodity personal computer (PC) [1,2] and hardware (such as field programmable gate array (FPGA)). Despite the favorable outcome, the display of hologram is a difficult problem that imposes, to a certain extent, a bottleneck to the practical realization of the holographic technology. One of the major problems is that high resolution devices that are capable of displaying holograms, such as Liquid Crystal on Silicon (LCOS), are only capable of displaying either the magnitude or the phase component of a complex hologram. A straightforward solution to this problem, is to employ two spatial light modulators (SLMs) to display the real and imaginary components, or the amplitude and phase information, of a complex hologram [3–5]. Likewise, a complex hologram can be simulated with a double phase-only hologram [6], and displayed with a pair of phase-only SLMs. In some implementation, the pair of SLMs is replaced by a single device, displaying a pair of holograms and subsequently merging the reconstructed wavefront through a grating [7–10]. Although such approach is effective, the optical setups are rather complicated. Besides, the area of the SLM allocated to each component of the hologram is reduced to half of its original size. Alternatively, a complex hologram can be converted into either an amplitude-only, or a phase-only hologram (POH) so that it can be displayed directly with a single device. However, if an amplitude-only display is used to display a Fresnel hologram, the reconstructed image will be contaminated with a de-focused twin image. While the real twin image can be diverted away from the reconstructed virtual image by converting an on-axis hologram into an off-axis hologram, the angular separation between the two images is limited by the resolution of the display. Besides, the optical efficiency of an amplitude hologram is rather low as the illumination beam is attenuated by the opacity of the fringe patterns. A POH, on the other hand, can be displayed with a phase-only device and results in higher optical efficiency as well as rejection of the twin image. On the downside, removing the magnitude component will lead to heavy distortion on the reconstructed image. To alleviate this problem, the Gerberg-Saxton algorithm [11], or iterative Fresnel transform [12] are often adopted to compute the phase hologram in an iterative manner, so that the reconstructed image will match with a target planar image. However, the phase hologram generated with such approach is computationally intensive. The fastest method of generating a POH developed to date, is probably the "One-Step Phase Retrieval (OSPR) process" [13,14]. Briefly, a random phase is first added to the object points prior to the generation of the digital hologram. The phase component of the hologram, which may be quantized with thresholding or error diffusion [15], is displayed with a phase-only device. The reconstructed images of the holograms generated with the OSPR process are generally noisy and multiple sub-frames, each representing the same object scene added with different random phase patterns, have to be presented rapidly to the observers to average out the speckle noise. However, the sub-frames involve more computation to generate, and also need to be displayed at high frame rate to avoid flickering. In this paper, we propose a novel and fast method for converting a complex Fresnel hologram into a single phase-only hologram called the bidirectional error

diffusion (BERD) hologram. The BERD hologram is capable of representing an object scene and preserving favorable visual quality on the reconstructed image. Different from the OSRP scheme, or other iterative approaches that generate a phase-only hologram from an object scene, our proposed method can be applied directly on a given digital complex hologram.

2. Bidirectional error diffusion (BERD) hologram as a novel phase-only hologram

2.1 Basic approach based on unidirectional error diffusion

To begin with, consider a complex Fresnel hologram that is generated from the object waves emitted from each point on an object scene. Each pixel in the hologram is denoted by $P_{u,v}$ and having a value $H(u,v)$ (where u and v are the vertical and horizontal axes of the coordinate system, respectively, as shown in Fig. 1(a)) given by

$$H(u,v) \Big|_{\substack{0 \leq u < X \\ 0 \leq v < Y}} = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \frac{I(x,y) \exp(i2\pi r_{x,y;u,v} / \lambda)}{r_{x,y;u,v}}, \quad (1)$$

where $I(x,y)$ and $r_{x,y;u,v} = [(u-x)^2 \delta + (v-y)^2 \delta + w_{x,y}^2]^{0.5}$ are the intensity of the point located at (x,y) in the object scene and its distance to the hologram, respectively. The perpendicular distance of a point at position (x,y) to the hologram is denoted by $w_{x,y}$, λ is the wavelength of the optical beam, and δ is the pixel size of the hologram. X and Y are the number of rows and columns of the hologram, and which are assumed to be the same as the object scene. We have assumed that no random phase has been added in the hologram generation process. The complex hologram can be converted into a POH, $H_p(u,v)$, by setting the magnitude of each pixel to be transparent with a value of unity, as given by

$$|H_p(u,v)| = 1, \text{ and } \arg(H_p(u,v)) = \arg(H(u,v)). \quad (2)$$

However, as we shall show later, the quality of the reconstructed image of a POH generated with Eq. (2) is extremely poor. Apparently, the heavy distortion on the reconstructed image is caused by the large amount of error in each hologram pixel after removing the magnitude information. In view of the above problem, we propose to employ the Floyd-Steinberg error diffusion technique [16] to compensate the error of each hologram pixel. The process is outlined as follows. To start with, each pixel in the hologram is scanned sequentially in a row by row, left to right manner, and the value of the pixel under evaluation is converted to a phase-only quantity $H_p(u,v)$ according to Eq. (2). Suppose P_{u_j,v_j} is the current pixel that is being processed. By setting the magnitude to unity will result in an error given by

$$E(u_j,v_j) = H(u_j,v_j) - H_p(u_j,v_j). \quad (3)$$

Next, the error $E(u_j,v_j)$ is diffused to the neighborhood pixels that have not been visited previously. Its neighborhood members are updated according to the following equations:

$$H(u_j,v_j+1) \leftarrow H(u_j,v_j+1) + w_1 E(u_j,v_j), \quad (4)$$

$$H(u_j+1,v_j-1) \leftarrow H(u_j+1,v_j-1) + w_2 E(u_j,v_j), \quad (5)$$

$$H(u_j+1,v_j) \leftarrow H(u_j+1,v_j) + w_3 E(u_j,v_j), \quad (6)$$

$$H(u_j+1, v_j+1) \leftarrow H(u_j+1, v_j+1) + w_4 E(u_j, v_j), \quad (7)$$

where the operator “ \leftarrow ” denotes updating the variable on the left hand side with the one on the right hand side of the expression. The values of the constant coefficients w_1 to w_4 are set to the values adopted in [16], with $w_1 = 7/16$, $w_2 = 3/16$, $w_3 = 5/16$, and $w_4 = 1/16$. According to Eqs. (4) to (7), we show, in Fig. 1(b), the spatial relation between pixel P_{u_j, v_j} and its neighborhood pixels for compensating the error via diffusion. The figure shows how P_{u_j, v_j} diffuses errors to its four unvisited neighborhood pixels. Equivalently, Eqs. (4-7) can be rearranged into a compact recursive expression given by

$$H(u_j, v_j) \leftarrow H(u_j, v_j) + w_1 E(u_j, v_j - 1) + w_4 E(u_j - 1, v_j + 1) + w_3 E(u_j - 1, v_j) + w_2 E(u_j - 1, v_j - 1) \quad (8)$$

and according to Eq. (8), Fig. 1(c) shows the updated value of each pixel can be visualized as the weighed sum of its four neighboring pixels. After all the hologram pixels have been visited and processed, the original hologram will be converted into a POH. As we shall illustrate in the experiment section, the reconstructed image of $H_p(u, v)$ generated with the proposed method is similar to the one obtained with the original complex hologram.

2.2 Enhanced method based on bidirectional error diffusion process

The basic approach described in section 2.1, though effective, results in noise contamination on the reconstructed image. A possible explanation is that in the error diffusion process, as given in Eq. (8), each hologram pixel is summed up with the errors diffused from neighboring pixels at fixed relative positions. This constitutes to a process which is similar to a predictive coder in which the error generated, to a certain extent, is correlated to the signal. As a result, the reconstructed image of the hologram, $H_p(u, v)$, will be superimposed with the noise signal corresponding to the correlated error. To overcome this problem, we propose an enhanced method to reduce the correlation between the holographic signal and the error. To begin with, we note that the error diffusion process represented in Eqs. (4-7) is independent on the direction of scanning the pixels along each row of the hologram image. As such, we change the scanning direction of each row of pixels so that the recursive process in Eq. (8), which leads to the correlated error, will be interrupted after each row is visited. The errors on the odd rows are diffused to the neighboring pixels in the same way as given in Eqs. (4-7), while the even rows are compensated by applying the error diffusion process in the opposite direction (from right to left), as given in Eqs. (9-12) below. The spatial relation between pixel P_{u_j, v_j} and its neighborhood pixels is illustrated in Fig. 1(d).

$$H(u_j, v_j - 1) \leftarrow H(u_j, v_j - 1) + w_1 E(u_j, v_j), \quad (9)$$

$$H(u_j + 1, v_j + 1) \leftarrow H(u_j + 1, v_j + 1) + w_2 E(u_j, v_j), \quad (10)$$

$$H(u_j + 1, v_j) \leftarrow H(u_j + 1, v_j) + w_3 E(u_j, v_j), \quad (11)$$

$$H(u_j + 1, v_j - 1) \leftarrow H(u_j + 1, v_j - 1) + w_4 E(u_j, v_j). \quad (12)$$

The resulting generated hologram based on Eqs. (4-7) to process the even rows and on Eqs. (9)-(12) to proceed the odd rows of the complex hologram is called the bidirectional error diffusion (BERD) hologram.

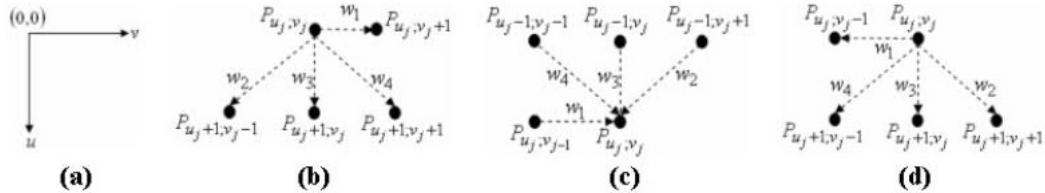


Fig. 1. (a) Horizontal and vertical axes of the co-ordinate system. (b) Diffusion of the error from the current pixel to its neighborhood, scanned from left to right. (c) An alternative representation of (a), showing the updating of a pixel from its neighborhood. (d) Diffusion of the error from the current pixel to its neighborhood, scanned from right to left.

3. Experimental results

A binary image “CITYU”, a smooth image “Lena”, and a highly textural image “Mandrill”, as shown in Figs. 2(a)-2(c), respectively are employed to evaluate our proposed method. Equation (1) is applied to generate the complex Fresnel holograms of the 3 images. The hologram, as well as the test images, are comprising of 2048×2048 pixels, each having a square size of $7\mu\text{m} \times 7\mu\text{m}$. The wavelength of the optical beam is 650nm . All the images are parallel to, and located at 0.3m from the hologram plane. The numerical reconstructed images of the 3 complex holograms are identical to the original images, and hence not shown in here. Next, we remove the magnitude component of the holograms with Eq. (2), and the numerical reconstructed images are shown in Figs. 3(a)-3(c). It can be seen that the shaded regions of the three images are removed extensively, leaving behind mostly the edges of the images. To overcome this problem, our proposed unidirectional error diffusion method described in section 2.1 is applied to convert the complex holograms into phase-only holograms (POHs). Reconstructed images of the POHs are shown in Figs. 4(a)-4(c). We observe that the reconstructed images are close to the original images in Figs. 2(a)-2(c). However, certain amount of noise contamination is noted, especially in the dark areas. Subsequently, we convert the complex holograms with our proposed bidirectional error diffusion method to obtain the BERD holograms as described in section 2.2. The numerical reconstructed images of the BERD holograms generated with such means, as shown in Figs. 5(a)-5(c), are very similar to the original images, and the noise signals are not noticeable. Quantitative evaluation on the quality of the reconstructed images of the POHs derived from the 3 aforementioned methods, as compared from the ones obtained with the complex holograms, are given in Table 1. Evidently, the proposed bidirectional error diffusion method generating BERD holograms results in significantly higher fidelity (in terms of PSNR) over its peers. For the sake of clarity, Figs. 2-5 are presented in their actual size in the media files MF2-MF5, respectively.

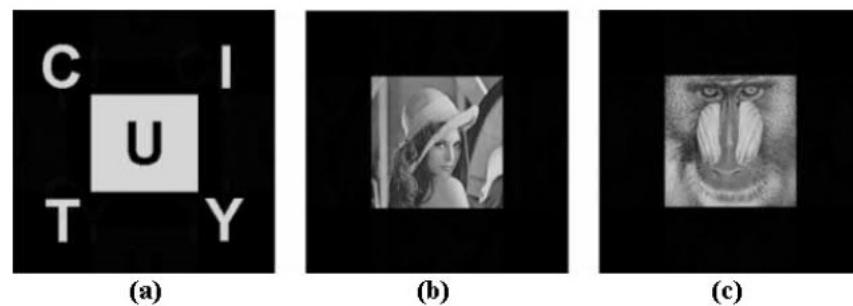


Fig. 2. (a)-(c) Test images “CITYU”, “Lena”, and “Mandrill”.

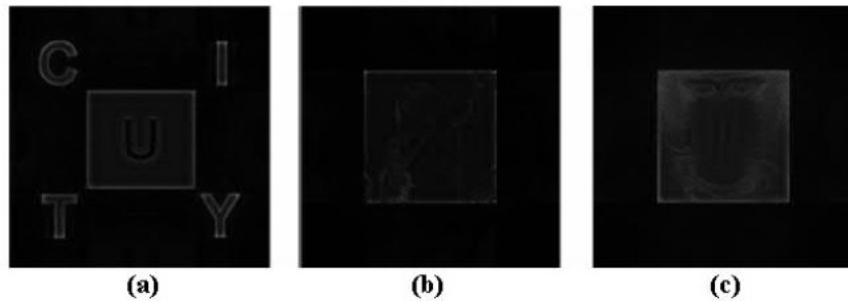


Fig. 3. (a)-(c) Reconstructed images of the phase-only holograms derived from direct removal of the magnitude component of the complex holograms.



Fig. 4. (a)-(c) Reconstructed images of the phase-only holograms derived from our proposed unidirectional error diffusion method.



Fig. 5. (a)-(c) Reconstructed images of the BERD holograms (novel phase-only holograms derived from our proposed bidirectional error diffusion method).

Table 1. Quantitative comparison between the fidelity of the reconstructed images as shown in the Figs. 3(a)-3(c), 4(a)-4(c), and 5(a)-5(c) (with reference to the ones obtained from the original complex holograms in Figs. 2(a)-2(c)) from phase-only holograms derived from different means. The reported values are PSNR values.

	CITYU	Lena	Mandrill
Direct removal of the magnitude component	13.0 dB	16.1 dB	16.2 dB
Unidirectional error diffusion	24.3 dB	28.1 dB	27.0 dB
Bidirectional error diffusion	31.6 dB	38.6 dB	32.7 dB

4. Conclusion

This paper reports a novel and fast method for converting a digital, complex Fresnel hologram into a phase-only hologram called the BERD holograms. The process is accomplished with a BERD mechanism, with which each hologram pixel is scanned sequentially and forced to a constant magnitude, while the resulting error is diffused to the unvisited neighborhood. The odd and even rows are scanned from opposite directions, constituting to a bidirectional error diffusion process. Our proposed method has 3 major advantages. First, the reconstructed images of the novel BERD holograms generated are very

similar to those obtained with the original complex holograms. Second, our proposed method only involves moderate amount of computation. Third, our method can be applied directly on an existing complex Fresnel hologram, without the presence of the object image. A potential research direction is to explore a better set of error diffusion coefficients (w_1 to w_4) that can enhance the quality of the reconstructed image of the BERD hologram.

Acknowledgments

This work is supported by the Chinese Academy of Sciences Visiting Professorships for Senior International Scientists Program under Grant Number 2010T2G17 and the High-End Foreign Experts Recruitment Program, China, under Grant Number GDJ20130491009.

附录 C 公式推导

附C1 魔法：物理只允许发生+数学却保证发生

附C1.1 物理：产生非线性极化波的能量源——偶极子

附C1.1.1 数学：若存在一个符合实验的表达式，则它必是张量的多级展开式

$$\begin{aligned}
 \tilde{\mathbf{P}}_n(\mathbf{r}, t) = \tilde{\mathbf{P}}(\omega_n) &= \varepsilon_0 \left[\sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}}} \overset{\rightarrow}{\chi}_e^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}(\omega_{m_{\pi(1)}}) + \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}}} D_{m_1, m_2} \overset{\rightarrow \rightarrow}{\chi}_e^{(2)}(\omega_n; \omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}) \right. \\
 &\quad : \tilde{\mathbf{E}}(\omega_{m_{\pi(1)}}) \tilde{\mathbf{E}}(\omega_{m_{\pi(2)}}) + \dots + \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}} + \dots + \omega_{m_{\pi(l)}}} D_{m_1, m_2, \dots, m_l} \overset{\overbrace{\rightarrow \rightarrow \dots \rightarrow}^{l+1}}{\chi}_e^{(l)}(\omega_n; \overbrace{\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}, \dots, \omega_{m_{\pi(l)}}}^l) \\
 &\quad \left. \dots \overbrace{\tilde{\mathbf{E}}(\omega_{m_{\pi(1)}}) \tilde{\mathbf{E}}(\omega_{m_{\pi(2)}}) \cdots \tilde{\mathbf{E}}(\omega_{m_{\pi(l)}})}^l + \dots \right] \tag{C1}
 \end{aligned}$$

其中，求和下限表遍历符合条件 $\omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}} = \omega_n$ 的单色光场组合 $(\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}})$ 的集合 $\{(\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}})\}$ （以 $l=2$ 为例），下标 n, m 对应的 $\{\omega_n\}, \{\omega_m\}$ 分别表示电极化强度 $\tilde{\mathbf{P}}$ 和电场强度 $\tilde{\mathbf{E}}$ 各自的频谱 ($\because n=m \not\Rightarrow \omega_n=\omega_m$ ，即下标字母的符号相异 \Leftrightarrow 不同频谱；同一频谱的 ω 的下标变量 n (或 m) 的取值只与相应频谱的 ω 的取值一一对应 (函数)^①)，即 $n_0 \rightarrow \omega_n|_{n=n_0}$ 、 $m_0 \rightarrow \omega_m|_{m=m_0}$ ；不同频谱的下标与本体无法跨频谱映射： $n_0 \not\rightarrow \omega_m|_{m=n_0}$ 、 $m_0 \not\rightarrow \omega_n|_{n=m_0}$ 。

并规定对 $\tilde{\mathbf{E}}$ 的频谱 $\{\omega_m\}$ 而言， $\{m\} \leftrightarrow \{\omega_m\}$ ；同理对 $\tilde{\mathbf{E}}$ 的单色光波集 $\{\overline{\omega_m}\}$ ，也有 $\{i\} \leftrightarrow \{\overline{m_i}\} \leftrightarrow \{\overline{\omega_{m_i}}\}$ ；且二者之间满足： $\{i\} \leftrightarrow \{\overline{m_i}\} \xrightarrow[\text{恒满射}]{\text{非单射}} \{m_i\} \leftrightarrow \{m\}$ ，即即使 $\{i\}$ 中 $i_1 \neq i_2$ ，也可以有 $m_{i_1} = m_{i_2}$ ，此时 m_{i_1}, m_{i_2} 在 $\{m_i\}$ 中值相同，简并为同一频率；在 $\{\overline{m_i}\}$ 中值不同，表频率相同的两束单色光，可见 $\{i\}$ 与 $\{m\}$ 不一一映射。说明 $\tilde{\mathbf{E}}$ 的 {单色光} $\xrightarrow[\text{满射}(\checkmark)]{\text{单射}(x)} \tilde{\mathbf{E}}$

-
- ① a. 映射. 属性=满(原像)对非空{像} \wedge 非一(原像)对多(像)=满对非空 \wedge 一一对应 (函数)；
 - b. 单射. 属性=映射. 属性 \wedge 非多对一=满对非空 \wedge 一对一；
 - c. 满射. 属性=映射. 属性 \wedge 非空对满=满对满 \wedge 非一对多；
 - d. 双射 \equiv 一对一映射 \equiv 一一映射 \equiv 一一对应 (映射) = 满射 \cap 单射；
 - I. 双射. 属性=单射. 属性 \wedge 满射. 属性=满对满 \wedge 一对一；
 - e. (单值)函数:=定义域 $\{x\}$ ($\subset \{\text{原像}\}$) 到值域 $\{y\}$ ($\subset \{\text{像}\}$) 的满射；
 - I. 函数. 属性=满射. 属性 \wedge {原像} & {像} 均是数集=满对满 \wedge 非一对多 $\wedge \{x\} \& \{y\}$ 是数集；

的{频率}，即频率相同的多束参与作用的 \tilde{E} 光，在 $\{\overline{\omega_{m_i}}\}$ 内仍是互不同的元素 ω_{m_i} ，但在 $\{\omega_m\}$ 内简并为同一 ω_m 。

附C1.1.2 数学符号的数学定义&物理含义

D_{m_1, m_2, \dots, m_l} 及各处数学符号的来源：考虑参与转换/合成的光场 $\{\tilde{E}_i\}$ 的频谱 $\{\omega_m\}$ 对第 l 项即第 l 阶非线性效应 $\tilde{P}^{(l)}(\omega_n)$ （单色而非频谱；无论通式或子项）的贡献：由于参与第 l 阶非线性过程的光电场 $\{\tilde{E}_i\}$ 的频率取值互相独立（可两两不等、可部分相等、亦可两两相等），所以 $\{\overline{\omega_{m_i}}\}$ 内可有多个光电场*i*的频率属性 ω_{m_i} 在频谱上均简并为同一个频率 ω_m ，这意味着 $\{\{i_{\omega_m}\}\}$ 中，任何一个一级元素 $\{i_{\omega_m}\}$ 内的二级元素 i_{ω_m} 的数量 $card\{i_{\omega_m}\}$ 无限制；所以参与任何非线性过程的光电场数目，即 $\{\overline{\omega_{m_i}}\} \leftrightarrow \{\{i_{\omega_m}\}\}$ 的最底层元素的数量，没有上限；也就是说，只要 $\{\{i_{\omega_m}\}\}$ 内有一个一级元素 $\{i_{\omega_m}\}$ ，那么由于其内元素的频率简并度可达 ∞ ，于是对于 $\{\overline{\omega_{m_i}}\}$ 或 $\{\{i_{\omega_m}\}\}$ 整体，其最底层元素数量 $card\{\overline{\omega_{m_i}}\} = card\{\{\overline{i_{\omega_m}}\}\} = card\{\{\overline{i_{\omega_m}}\}\} \times card\{i_{\omega_m}\}$ ，上限为 $card\{\{\overline{i_{\omega_m}}\}\} \times \infty$ ，其中 $card\{\{\overline{i_{\omega_m}}\}\}$ 一般为有限值。

附C1.1.2.1 对各物理量本体与各级角标的物理意义、取值数目和取值上限的规范

虽对*i*的取值数量没有要求，但由于每个*i*代表一个参与转换的 \tilde{E}_i ，而除非所有光学参数相同，否则两束光不是同一束光，则要求*i*两两相异（数学上由于 $\{i\}$ 是个集合，也要求其内元素两两相异）；因此对每个*i*，没有其他取值要求，以及取值范围的要求，只需两两取值相异——但为确定起见，这里规定，*i*从1开始，按照 N^+ 集合内元素值的大小递增来标号，即*i*=1,2,3,...，因此有：

$$\textcircled{1} = i_{max} = card\{i\} = \dots = card\{\overline{m}\} = card\{\overline{\omega_m}\} = \dots = card\{\tilde{E}_i\} = l \in N^+ \quad (\text{C2})$$

①值的大小由所考虑的非线性光学效应的阶数决定（当然，能^①参与混频的过程的单色入射光数目L也得 $\geq l$ ，这样才有能力提供*l*束单色光来实际参与混频；这是很容易满足的条件，只要能精确地满足能量守恒 $\omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}} + \dots + \omega_{m_{\pi(l)}} = \omega_n$ ，连单束入射单色光都可做到，因为上文说过 $card\{i_{\omega_m}\}$ 上限是 ∞ ）；且相对于②独立取值，上不封顶。

^① 对数学表达式C1，接下来只关心参与混频的*l*束单色入射光，不关心除此以外的，能参与但没参与的光及其频率；但式C1是通式，因而包含了能参与但没参与*l*阶非线性效应的那*L-l*束光。

由于实际可参与混频的单色入射光数目可能会多于所考虑的非线性光学效应的阶数即 $L \geq l$ ，式 C2 的适用范围为考虑某一具体的非线性过程，如考虑给定的 l 束光的 l 阶非线性过程，此时参与混频的单色场数目必须与非线性过程的阶数相同，不能多也不能少。而对于式(C1)所描绘的一般情况，即若考虑 $L \geq l$ 束混频光的 l 阶非线性效应，则 $\textcircled{1}=i_{max} = L$ ，且光束序列(号的排列) π 的长度也将从 l 扩展为 L ，下面的 $\textcircled{2} \leq l$ 也得修正为 $\textcircled{2} \leq L$ ，即总有 $\textcircled{2} \leq \textcircled{1}$ ，之后的③中求和上限 m_{max} 也得进一步修正，给接下来的描述带来很多不便；且由于即使考虑的光场对象与非线性阶数对象的数目满足上述关系，且非线性过程数目不止一个而很多 ($\leq A_L^l = L!/l!$)，在考虑某一具体的 l 阶非线性过程的时候，总是 l 束光而非 L 束光参与的，因此总可以对所考虑的非线性过程中参与混频的光重新编排角标为 $1 \sim l$ ，且将相应的频谱和频率角标重新排序，以至于这里以及接下来的结论仍成立。

同样由于频谱 $\{\omega_m\}$ 内两两频率 ω_m 取值不同，而频率下标取值集合 $\{m\}$ 与频率取值集合 $\{\omega_m\}$ 一一对应，所以频率下标 m 也两两不同（数学上由于 $\{m\}$ 是个集合，也要求其内元素两两相异）；为确定起见，仍规定 $m=1,2,3,\dots$ （对 m 和对 i 的规定，适用于诸如 $\overline{\{m_i\}}$ 、 $\overline{\{\omega_m\}}$ 、 $\{m_i\}$ 、 $\{\omega_{m_i}\}$ 等一切含 m 或含 i 的集合或元素），因此有：

$$\textcircled{2}=m_{max}=card\{m\}=card\{\omega_m\} \leq l \quad (\text{C3})$$

②值的大小由参与合成的 \tilde{E}_i 的频率属性，所构成的集合，即参与合成的 $\{\tilde{E}_i\}$ 的频谱决定，即等于其离散谱的峰的数量；虽②的取值上限受①的取值限制，但②的取值本身也与①的取值无关，即取值相互独立（即不构成函数关系）；

利用频率因作为光束的标签/属性之一而可以简并（到频谱中去）的思想，可将能量守恒条件 $\omega_{m_1} + \omega_{m_2} + \dots + \omega_{m_l} = \omega_n$ 合并同类项，从纯加和的形式转而写为“一名 1850s 年代研究组合数学的数学家 Francesco Faà di Bruno 所给出的复合函数高阶导通式(Faà di Bruno's formula)的，每项所满足的条件，以及为求子项总数所需遍历的方程，即 $b_1 + 2b_2 + \dots + mb_m = m$ ，且 $b_1 + b_2 + \dots + b_m = k$ ”^①的先乘后求和的形式：

$$\sum_m card\{i_m\} \times \omega_m = \omega_n, \text{ 且 } \sum_m card\{i_m\} = l, \text{ 即:}$$

^① 为对应，这里将原本的 $k := b_1 + b_2 + \dots + b_m$ 写作 $b_1 + b_2 + \dots + b_m = k$ ，注意其本义是在定义 k 而非约束条件。

$$\textcircled{3}. \sum_{m=1}^{m_{max}} \text{card}\{i_m\} \times \omega_m = \omega_n, \text{ 且 } \sum_{m=1}^{m_{max}} \text{card}\{i_m\} = l \quad (\text{C4})$$

③.是两个必须同时满足的条件，不是一个表达式，也就没有其值的说法，这是其在类型上与①②的主要区别（③.右下角含点，①②不含）；但③.又是连接起①②的桥梁，特别是第二个条件，给出了①与②的关系式，即 m_{max} 与 $l=i_{max}$ 间的关系，二者的取值差异由参与混频的光束集 $\{i\}$ 关于参与混频的各频率 m 的简并度 $\text{card}\{i_m\}$ 决定——某单个频率 m_k 对应的简并度 $\text{card}\{i_{m_k}\}$ 越大，或简并度 $\text{card}\{i_{m_k}\} \geq 2$ 的频率 m_k ^① 越多，则 m_{max} 与 l 的差距 $(l - m_{max})/l$ 越大。

从③.可知，相对于所考虑的 ω_n ，即使参与混频的入射光场是单色弱频的，即频谱 $\{\omega_m\}$ 内只有一个峰及其横坐标 ω_m ， $\text{card}\{\omega_m\}=1$ ，且单独的离散峰横坐标 ω_l 小于所考虑的 $\tilde{P}^{(l)}(\omega_n)$ 中的 ω_n ，仍可能存在二阶及以上的高阶非线性光学效应，此时虽然 $m_{max}=1$ ，但 l 可以取适当值，以使得恰有：

$$\begin{aligned} \sum_{m=1}^{m_{max}} \text{card}\{i_m\} \times \omega_m &\xrightarrow{\text{③. 中一式: 能量守恒}} \text{card}\{i_1\} \times \omega_1 \\ &\xleftarrow{\text{③. 中二式: 数量守恒}} \sum_{m=1}^{m_{max}} \text{card}\{i_m\} \times \omega_1 = l \times \omega_1 = \omega_n \end{aligned} \quad (\text{C5})$$

若反客为主地以频率为 ω_l 的参与合成的单色入射光为考察对象（不论之前考察某具体 ω_n 还是这里考察 $\omega_m|_{m=1}$ ，均不关心参与合成的同频率 ω_l 的单色光数目），考察其可能参与的所有非线性过程/混频过程，则它将不仅可能，并且一定（如果其他条件允许），同时参与 $l=2,3,4,\dots$ 的所有高次极化波的合成过程，分别对应 $\text{card}\{2,3,4,\dots\} = \text{card}\{N^+\} - 1$ 个式 1-1 的第 2,3,4,... 阶非线性光学效应，并同时合成出光束数目 $\text{card}\{j\} = \text{频谱}\{\omega_n\} = \{l \times \omega_1\}$ 内频率数 $\text{card}\{\omega_n\}$ 的，相应的倍频极化光集 $\{j\}$ ；只是功率谱上频率越高的倍频极化光，功率越低（因非线性光学过程转换效率随阶数的增高而降低），其产生的高次谐波的功率也低，而不容易被仪器探测到，并且也是能量守恒定律的要求：高阶非线性过程产生的高频极化光（子）的频率高，则能量高，因此高频极化光（子）

^① 用 m_k 表示不关心该频率是属于哪条光束，即 m_k 是不关心频率的光束属性的（ m_i 就是要关心其属于哪条光束的；单独提到 m_i 一般认为是 $\overline{\{m_i\}}$ 而非 $\{m_i\}$ 集合内的元素，但这里与 m_k 作比较的 m_i 是 $\{m_i\}$ 内的元素）、参与混频的频谱中的、某个具体频率 m （ k 与 m 的取值范围相同），正如不关心光束的频率属性的某个具体光束 i 可以用 i_k 表示一样（若与频率牵连/链接则用 i_m ，它是集合 $\{i_m\}$ 中的元素；此处 k 与 i 的取值范围相同）；而 m_{i_k} 即表示具体某频率未知的光束 i_k 的频率，其下标 i_k 本体是 i ，所以与 m_i 联系颇甚，但要注意它与 m_k 的区别甚大。

得产生得少，才可能符合能量守恒。而这一过程也可以由各阶电极化率系数(的各分量大小)的调整而得到保障。

附C1.1.2.2 非线性过程所合成的极化光一般与其频率一一对应，而合成原料并不一定

另需要解释一下，标量形式(2.21)和张量形式(2.1)的非线性表达式，由于在物理意义上是光束上的而非频率上的“多 $\tilde{\mathbf{E}} \rightarrow$ 单 $\tilde{\mathbf{P}}$ ”，因此除了 $\tilde{\mathbf{E}}$ 阵营的有关各量下标必须含 i 而不必含 m 之外，按理说 $\tilde{\mathbf{P}}$ 阵营的有关各量下标也必须含 j 而不必含 n ，但在这些写法中， $\tilde{\mathbf{P}}$ 阵营的有关各量却只含了 n ，未含 j ——这是因为尽管理论上，单看 $\tilde{\mathbf{P}}$ 阵营的话， j 不同的光束，频率及 n 可能相同，但在实验上，或者单纯从方便考虑的角度，由 $\tilde{\mathbf{E}}$ 所合成的 $\tilde{\mathbf{P}}$ ， j 不同则 n 不同，不论 $\{\tilde{\mathbf{P}}_j\}$ 是由“多 $\tilde{\mathbf{E}} \rightarrow$ 单 $\tilde{\mathbf{P}}$ ”还是由“多 $\tilde{\mathbf{E}} \rightarrow$ 多 $\tilde{\mathbf{P}}$ ”过程合成的，因为对于指定 ω_n ，各阶非线性过程所合成的 $\tilde{\mathbf{P}}^{(l)}(\omega_n)$ ，都放入同一个 $\tilde{\mathbf{P}}(\omega_n)$ 中了；除非构造特殊的实验场景或物理情景，才有可能对于不同的光束 j ，有相同的频率 n ，或考虑出射光而不是合成光；而 $\tilde{\mathbf{E}}$ 阵营的有关各量下标不能不含 i ，便是因为不同 i 的 $\tilde{\mathbf{E}}$ 光的频率 m 可能相同。

而之所以 $\tilde{\mathbf{E}}$ 阵营的有关各量的下标除了 i 之外，还含 m ，是为了方便之后引入简并因子 D 中的 $card\{\overline{m_i | m_i = m}\}$ 等概念，同时也是避免 $\tilde{\mathbf{E}}_i$ 被误认为 $\tilde{\mathbf{E}}_m|_{m=1}$ 而非 $\tilde{\mathbf{E}}_i|_{i=1}$ ，而 $\tilde{\mathbf{E}}_{m_i}$ 就可避此嫌。

附C1.2 非线性电介质中的波动方程

附C1.2.1 从麦氏方程组的微分形式到非线性电介质中的波动方程

$$\left\{ \begin{array}{l} \nabla \cdot \tilde{\mathbf{D}} = \tilde{\rho}_f \\ \nabla \cdot \tilde{\mathbf{B}} = 0 \\ \nabla \times \tilde{\mathbf{E}} = -\frac{\partial \tilde{\mathbf{B}}}{\partial t} \\ \nabla \times \tilde{\mathbf{H}} = \tilde{\mathbf{J}}_f + \frac{\partial \tilde{\mathbf{D}}}{\partial t} \end{array} \right. \quad (C6)$$

其中，物理量头上的睂化符“~”表示相应物理量是含时的(时变的)，与该物理量是标量、矢量还是张量无关，也与物理量是否随空间坐标变化无关；且 $\tilde{\mathbf{D}}, \tilde{\mathbf{B}}, \tilde{\mathbf{E}}, \tilde{\mathbf{H}}$ 都是叠加场，非单色场，比如 $\tilde{\mathbf{D}}$ 中的 $\tilde{\mathbf{P}}$ ，虽在“多 $\tilde{\mathbf{E}} \rightarrow$ 单 $\tilde{\mathbf{P}}$ ”的便于考虑的理想化环境下满足 $\tilde{\mathbf{P}}(\omega_n) = \sum_{l=1}^{\infty} \tilde{\mathbf{P}}^{(l)}(\omega_n)$ ，且仍属于单色，但实际总会发生“单 $\tilde{\mathbf{E}} \rightarrow$ 多 $\tilde{\mathbf{P}}$ ”的倍频过程(如果其他条件满足)，此时单一频率 ω_n 变成频谱 $\{\omega_n\}$ ，衍生出多个频率不同

的高次极化波，合并在 $\tilde{\mathbf{P}}$ 中，因此实际情况下 $\tilde{\mathbf{P}}$ 不是单色的；所以即使麦氏方程组→波动方程→非线性极化波→和频波，这一过程的落脚点是某一特解单色波 $\tilde{\mathbf{E}}_n$ ，初始的麦氏方程组中包含通解的 $\tilde{\mathbf{D}}, \tilde{\mathbf{B}}, \tilde{\mathbf{E}}, \tilde{\mathbf{H}}$ 各量，都是复合场，尽管其也适用于单色场。

由麦氏方程组所得的非线性介质中电场波动方程、磁场波动方程的最普遍形式：

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}) + \frac{1}{c^2} \frac{\partial^2 \tilde{\mathbf{E}}}{\partial t^2} = -\frac{1}{\epsilon_0 c^2} \left[\frac{\partial^2 \tilde{\mathbf{P}}}{\partial t^2} + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}) \right] \quad (C7)$$

$$\nabla \times (\nabla \times \tilde{\mathbf{H}}) + \frac{1}{c^2} \frac{\partial^2 \tilde{\mathbf{H}}}{\partial t^2} = -\frac{1}{c^2} \frac{\partial^2 \tilde{\mathbf{M}}}{\partial t^2} + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{P}}) \quad (C8)$$

两个波动方程之所以是普适的，离不开 $\tilde{\mathbf{H}} = \tilde{\mathbf{B}}/\mu_0 - \tilde{\mathbf{M}}$ 与 $\tilde{\mathbf{D}} = \epsilon_0 \tilde{\mathbf{E}} + \tilde{\mathbf{P}}$ 是普适的，因为二者是 Maxwell 方程组推导过程中定义的物理量。上述两个方程还可分别写作另外两种形式：

$$\begin{aligned} \nabla \times (\nabla \times \tilde{\mathbf{E}}) + \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \tilde{\mathbf{D}}}{\partial t^2} &= -\frac{1}{\epsilon_0 c^2} \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}) \\ \nabla \times (\nabla \times \tilde{\mathbf{H}}) + \frac{1}{\mu_0 c^2} \frac{\partial^2 \tilde{\mathbf{B}}}{\partial t^2} &= \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{P}}) \end{aligned} \quad (C9)$$

$$\begin{aligned} \nabla \times (\nabla \times \tilde{\mathbf{D}}) + \frac{1}{c^2} \frac{\partial^2 \tilde{\mathbf{D}}}{\partial t^2} &= \nabla \times (\nabla \times \tilde{\mathbf{P}}) - \frac{1}{c^2} \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}) \\ \nabla \times (\nabla \times \tilde{\mathbf{B}}) + \frac{1}{c^2} \frac{\partial^2 \tilde{\mathbf{B}}}{\partial t^2} &= \mu_0 \left[\nabla \times (\nabla \times \tilde{\mathbf{M}}) + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{P}}) \right] \end{aligned} \quad (C10)$$

但之所以不写成这两种形式，是因为这样写非线性量太多不便于求解。

并且接下来也主要使用式(C7)即电场波动方程，因为我们主要关心的是光电场，而维纳实验证实了电磁波对人眼和仪器起主要作用的是其电场部分；或者考察电磁波的 $|\mathbf{E}_n(\mathbf{r})|/|\mathbf{B}_n(\mathbf{r})|$ 比值，连同用原子物理的方法，可估算出光波的 $\tilde{\mathbf{E}}_n$ 对电子、介质宏观电偶极矩产生的力 $\rho_f dV \cdot \tilde{\mathbf{E}}_n$ 、 $(\tilde{\mathbf{P}} \cdot \nabla) \tilde{\mathbf{E}}_n = \tilde{\mathbf{P}} \cdot \nabla \tilde{\mathbf{E}}_n$ ，远大于光波的 $\tilde{\mathbf{B}}_n$ 对介质宏观磁化强度的作用力 $(\tilde{\mathbf{M}} \cdot \nabla) \tilde{\mathbf{B}}_n = \tilde{\mathbf{M}} \cdot \nabla \tilde{\mathbf{B}}_n$ 。

附C1.2.2 单色化 $\tilde{\mathbf{E}}, \tilde{\mathbf{P}}, \tilde{\mathbf{D}}$ 、引入 $\tilde{\mathbf{P}}, \tilde{\mathbf{D}}$ 的非线性效应

附C1.2.2.1 将 $\tilde{\mathbf{E}}$ 相关的线性项全都放于左侧，非线性项放在右侧作为零次常系数部分

暂忽略 $\nabla \times \tilde{\mathbf{M}}$ 所在项，则 $\tilde{\mathbf{D}}$ 中的非 $\tilde{\mathbf{E}}$ 部分(即 $\tilde{\mathbf{P}}$)或 $\tilde{\mathbf{D}}$ 的非线性部分(即 $\tilde{\mathbf{D}}^{NL} = \tilde{\mathbf{P}}^{NL}$ $= \sum_{l=2}^{\infty} \tilde{\mathbf{P}}^{(l)}$)，被认作为产生和频波的波源，或非线性波源，因无法分离出待求时变矢量 $\tilde{\mathbf{E}}$ ，而作为常系数项，放在方程右侧；而 $\tilde{\mathbf{D}}$ 中含 $\tilde{\mathbf{E}}$ 的部分，或 $\tilde{\mathbf{D}}$ 中的线性部分

$\tilde{\mathbf{D}}^{(1)} = \varepsilon_0 \tilde{\mathbf{E}} + \tilde{\mathbf{P}}^{(1)}$, 放在方程左侧, 因为这样一来可分离的显式待求时变矢量 $\tilde{\mathbf{E}}$ 全都放在方程左侧了 ($\tilde{\mathbf{P}}^{NL}$ 中也含 $\tilde{\mathbf{E}}$, 但含的是 $\tilde{\mathbf{E}}$ 的高次方, 且含有 $\tilde{\mathbf{E}}$ 的各分量; 而 $\tilde{\mathbf{P}}^{(1)}$ 虽也是 $\tilde{\mathbf{E}}$ 的各分量的合成, 但不含 $\tilde{\mathbf{E}}$ 的高次方, 因而在左边可以分离出)。

现将 $\tilde{\mathbf{D}} = \tilde{\mathbf{D}}^{(1)} + \tilde{\mathbf{D}}^{NL} = \tilde{\mathbf{D}}^{(1)} + \tilde{\mathbf{P}}^{NL}$ 代入式(C9)中的一式, 或将 $\tilde{\mathbf{D}}^{(1)} = \varepsilon_0 \tilde{\mathbf{E}} + \tilde{\mathbf{P}}^{(1)}$ 和 $\tilde{\mathbf{P}} = \tilde{\mathbf{P}}^{(1)} + \tilde{\mathbf{P}}^{NL}$ 代入式 1-10 中一式, 即得:

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}) + \frac{1}{\varepsilon_0 c^2} \frac{\partial^2 \tilde{\mathbf{D}}^{(1)}}{\partial t^2} = -\frac{1}{\varepsilon_0 c^2} \frac{\partial^2 \tilde{\mathbf{P}}^{NL}}{\partial t^2} - \frac{1}{\varepsilon_0 c^2} \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}) \quad (C11)$$

在各向异性的电介质中, $\tilde{\mathbf{P}}^{(1)} = \varepsilon_0 \vec{\chi}_e^{(1)} \cdot \tilde{\mathbf{E}}$ 中的 $\vec{\chi}_e^{(1)}$ 的主对角元之外的元素并非全为零, 所以没法退化为零阶张量即标量 $\chi_e^{(1)}$, 因此将

$$\tilde{\mathbf{D}}^{(1)} = \varepsilon_0 \tilde{\mathbf{E}} + \tilde{\mathbf{P}}^{(1)} = \varepsilon_0 \tilde{\mathbf{E}} + \varepsilon_0 \vec{\chi}_e^{(1)} \cdot \tilde{\mathbf{E}} = \varepsilon_0 (\vec{\mathbf{I}} + \vec{\chi}_e^{(1)}) \cdot \tilde{\mathbf{E}} = \varepsilon_0 \vec{\boldsymbol{\epsilon}}_r^{(1)} \cdot \tilde{\mathbf{E}} = \vec{\boldsymbol{\epsilon}}^{(1)} \cdot \tilde{\mathbf{E}} \quad (C12)$$

代入即有:

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}) + \frac{1}{c^2} \frac{\partial^2 (\vec{\boldsymbol{\epsilon}}_r^{(1)} \cdot \tilde{\mathbf{E}})}{\partial t^2} = -\frac{1}{\varepsilon_0 c^2} \frac{\partial^2 \tilde{\mathbf{P}}^{NL}}{\partial t^2} - \frac{1}{\varepsilon_0 c^2} \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}) \quad (C13)$$

在 $\vec{\boldsymbol{\epsilon}}_r^{(1)}$ 不含时的条件下, $\vec{\boldsymbol{\epsilon}}_r^{(1)}$ 可从第二项的偏导数的分子中提取出来:

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}) + \frac{1}{c^2} \vec{\boldsymbol{\epsilon}}_r^{(1)} \cdot \frac{\partial^2 \tilde{\mathbf{E}}}{\partial t^2} = -\frac{1}{\varepsilon_0 c^2} \frac{\partial^2 \tilde{\mathbf{P}}^{NL}}{\partial t^2} - \frac{1}{\varepsilon_0 c^2} \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}) \quad (C14)$$

该波动方程相对于式(C7)的电场波动方程, 即进化到了将 $\tilde{\mathbf{E}}$ 相关的线性项全都放在了左侧, 非线性项放在右侧, 而不是简单地仅将待求的 $\tilde{\mathbf{E}}$ 与 $\tilde{\mathbf{P}}$ 分居两侧。

附C1.2.2.2 引入非线性效应之前, 必须先单色化各场

但现在有两个问题, 一是式(C14)中各矢量都是复合场的话, 如何求解。既然麦氏方程组是普适的, 那么其以及其中各量也将适用于单色的情况, 这一方面允许我们单色化波动方程和其中各量; 二是以上过程中, $\tilde{\mathbf{D}}$ 和 $\tilde{\mathbf{P}}$ 的线性部分 $\tilde{\mathbf{D}}^{(1)} = \varepsilon_0 \tilde{\mathbf{E}} + \tilde{\mathbf{P}}^{(1)}$, $\tilde{\mathbf{P}}^{(1)}$, 以及 $\tilde{\mathbf{D}}$ 和 $\tilde{\mathbf{P}}$ 的非线性部分 $\tilde{\mathbf{D}}^{NL} = \tilde{\mathbf{P}}^{NL} = \sum_{l=2}^{\infty} \tilde{\mathbf{P}}^{(l)}$, 必须指定一个待合成的频率 ω_n , 因为 $\tilde{\mathbf{D}}$ 的各部分中均有 $\tilde{\mathbf{P}}$ 的部分, 而 $\tilde{\mathbf{P}}$ 的线性和非线性部分, 由式 1-1 给定, 而式 1-1 作为“多 $\tilde{\mathbf{E}} \rightarrow$ 单 $\tilde{\mathbf{P}}$ ”的合成过程, 必需存在一个待合成的 ω_n 以供考虑, 并且按式 1-1 只能合成出一个单色 $\tilde{\mathbf{P}}_n$ 。因此只要涉及“线性部分”或“非线性部分”, 则其主语 $\tilde{\mathbf{P}}$ 或 $\tilde{\mathbf{D}}$ 必须是单色的, 并且是由“多 $\tilde{\mathbf{E}} \rightarrow$ 单 $\tilde{\mathbf{P}}$ or 单 $\tilde{\mathbf{D}}$ ”过程生成的, 因此上述从 1-10 之后, 方程左右两侧区分了 $\tilde{\mathbf{E}}$ 的线性和非线性部分之后, 所有的 $\tilde{\mathbf{P}}, \tilde{\mathbf{D}}$, 非线性系数 (极化率张量) $\vec{\chi}_e$, 介电常数 (电容率) $\vec{\epsilon}$, 相对介电常数 $\vec{\epsilon}_r$, 包括其各阶非线性成

分如 $\tilde{\mathbf{P}}^{(1)}, \tilde{\mathbf{P}}^{NL}, \tilde{\mathbf{D}}^{(1)}, \tilde{\mathbf{D}}^{NL}$ 及 $\vec{\chi}_e^{(l)}, \vec{\varepsilon}_r^{(l)}, \vec{\varepsilon}_r^{(l)}$ ，其下标均应加上 n 。

而对于所有的 $\tilde{\mathbf{E}}$ ，按理说下标需加上 i 而不是 n ，因为参与合成的入射光集 $\{\tilde{\mathbf{E}}_i\}$ 中并不实际存在 $\tilde{\mathbf{E}}_n$ ；但别忘了就是要求该方程的解 $\tilde{\mathbf{E}}_n$ ，它本身就未知所以才需要求解，因此这里下角标可以用 n ，而且 $\tilde{\mathbf{E}}_n$ 的物理含义就是经非线性过程所合成的可出射晶体的和频光（非和频极化光）。它没有阶数可言，因为方程右侧产生各阶非线性效应的各阶非线性波源，全都在按能量守恒方程，通过 $l+1$ 波混频，将能量传输给此存在于 $\{\tilde{\mathbf{E}}_i\}$ 之外的第 $l+1$ 个参与混频的合成波 $\tilde{\mathbf{E}}_{l+1}$ ，作为最终的 $\tilde{\mathbf{E}}_n$ 的一部分；同时，为所有 $\tilde{\mathbf{E}}$ 加上下标 n 而不是 i ，这在数学上也是为了使式 1-12 的单色版能够两边同时消去共同的时间项相位因子便于求解。因此这就要求单色化波动方程和相关各物理量：

$$\begin{aligned} \tilde{\mathbf{D}}_n^{(1)} &= \varepsilon_0 \tilde{\mathbf{E}}_n + \tilde{\mathbf{P}}_n^{(1)} = \varepsilon_0 \tilde{\mathbf{E}}_n + \varepsilon_0 \vec{\chi}_e^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}_n = \varepsilon_0 [\vec{\mathbf{I}} + \vec{\chi}_e^{(1)}(\omega_n)] \cdot \tilde{\mathbf{E}}_n \\ &= \varepsilon_0 \vec{\varepsilon}_r^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}_n = \vec{\varepsilon}^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}_n \quad \textcircled{1} \end{aligned} \quad (\text{C15})$$

$$\tilde{\mathbf{D}}_n^{(l)} = \tilde{\mathbf{P}}_n^{(l)} (l \geq 2) \quad , \quad \tilde{\mathbf{D}}_n^{NL} = \tilde{\mathbf{P}}_n^{NL} = \sum_{l=2}^{\infty} \tilde{\mathbf{P}}_n^{(l)} \quad , \quad \tilde{\mathbf{D}}_n = \tilde{\mathbf{D}}_n^{(1)} + \tilde{\mathbf{D}}_n^{NL} \quad , \quad \tilde{\mathbf{P}}_n = \tilde{\mathbf{P}}_n^{(1)} + \tilde{\mathbf{P}}_n^{NL} \quad (\text{C16})$$

同时方程(C14)变为：

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}_n) + \frac{1}{c^2} \vec{\varepsilon}_r^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{E}}_n}{\partial t^2} = -\frac{1}{\varepsilon_0 c^2} \frac{\partial^2 \tilde{\mathbf{P}}_n^{NL}}{\partial t^2} - \frac{1}{\varepsilon_0 c^2} \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n) \quad (\text{C17})$$

附C1.2.3 单色化 $\tilde{\mathbf{H}}, \tilde{\mathbf{M}}, \tilde{\mathbf{B}}$ 、引入 $\tilde{\mathbf{M}}, \tilde{\mathbf{B}}$ 的非线性效应

附C1.2.3.1 波动方程右侧 $\tilde{\mathbf{M}}_n$ 也需要单色化和引入非线性项

同理，为了消去时间项，方程(C17)中的 $\tilde{\mathbf{M}}_n$ 也必须是单色且与其余量同频率；现在为进一步简化(C17)，需引入 $\tilde{\mathbf{M}}, \tilde{\mathbf{B}}$ 的非线性效应，而只要涉及“线性部分”或“非线性部分”，则其主语 $\tilde{\mathbf{M}}$ 或 $\tilde{\mathbf{B}}$ 也必须是单色的，并且是由“多 $\tilde{\mathbf{H}} \rightarrow$ 单 $\tilde{\mathbf{M}}$ or 单 $\tilde{\mathbf{B}}$ ”过程生成的，那么这也要求 $\tilde{\mathbf{H}}, \tilde{\mathbf{M}}, \tilde{\mathbf{B}}$ 均为单色。类似式(C1)地，非线性磁效应可以表示

^① $\tilde{\mathbf{D}}_n = \vec{\varepsilon}(\omega_n) \cdot \tilde{\mathbf{E}}_n$ 不成立， $\tilde{\mathbf{D}} = \vec{\varepsilon} \cdot \tilde{\mathbf{E}}$ 便更不成立；只能仿 $\tilde{\mathbf{P}}_n = \tilde{\mathbf{P}}_n^{(1)} + \tilde{\mathbf{P}}_n^{NL}$ 那样，写作 $\tilde{\mathbf{D}}_n$ 关于 $\tilde{\mathbf{E}}_n$ 的展开式 $\tilde{\mathbf{D}}_n = \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}}} \vec{\varepsilon}^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}(\omega_{m_{\pi(1)}}) + \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}}} D_{m_1, m_2} \vec{\varepsilon}^{(2)}(\omega_n) : \tilde{\mathbf{E}}(\omega_{m_{\pi(1)}}) \tilde{\mathbf{E}}(\omega_{m_{\pi(2)}}) + \dots$

为：

$$\begin{aligned}
 \tilde{\mathbf{M}}_n(\mathbf{r}, t) = \tilde{\mathbf{M}}(\omega_n) &= \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}}} \xrightarrow{\rightarrow} \chi_m^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}(\omega_{m_{\pi(1)}}) + \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}}} D_{m_1, m_2} \xrightarrow{\rightarrow \rightarrow} \chi_m^{(2)}(\omega_n; \omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}) \\
 &\quad : \tilde{\mathbf{H}}(\omega_{m_{\pi(1)}}) \tilde{\mathbf{H}}(\omega_{m_{\pi(2)}}) + \dots + \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}} + \dots + \omega_{m_{\pi(l)}}} D_{m_1, m_2, \dots, m_l} \xrightarrow{\overbrace{\rightarrow \rightarrow \dots \rightarrow}^{l+1}} \chi_m^{(l)}(\omega_n; \overbrace{\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}, \dots, \omega_{m_{\pi(l)}}}^l) \\
 &\quad \vdots \tilde{\mathbf{H}}(\omega_{m_{\pi(1)}}) \tilde{\mathbf{H}}(\omega_{m_{\pi(2)}}) \cdots \tilde{\mathbf{H}}(\omega_{m_{\pi(l)}}) + \dots
 \end{aligned} \tag{C18}$$

其中所有 ω_{m_i} 均与式(C1)中的 ω_{m_i} 完全相同，均是同一条光束 i 的角频率；相应地，所有以 ω_{m_i} 为自变量的因变量，也全都属于同一条光束 i ，比如式(C18)中的 $\tilde{\mathbf{H}}(\omega_{m_i})$ 与同 i 的式(C1)中的 $\tilde{\mathbf{E}}(\omega_{m_i})$ ，也均是同一条光束 i 的磁场和电场两条属性；同理，该式中的 $\tilde{\mathbf{M}}(\omega_n)$ 在生成的同时，也必伴随着式(C1)中的 $\tilde{\mathbf{P}}(\omega_n)$ 同时生成(在其光路附近)，反之亦然；可见上式中各阶非线性效应中参与混频的磁场(频率)序列 $\{(\omega_{m_{\pi(1)}}, \omega_{m_{\pi(2)}}, \dots, \omega_{m_{\pi(l)}})\}$ 、简并因子，也与式(C1)中相应阶磁场(频率)序列、简并因子分别相同。

类似关系(C15)、(C16)地，可将 $\tilde{\mathbf{M}}, \tilde{\mathbf{B}}$ 分别分为线性部分和非线性部分：

$$\begin{aligned}
 \tilde{\mathbf{B}}_n^{(1)} &= \mu_0 [\tilde{\mathbf{H}}_n + \tilde{\mathbf{M}}_n^{(1)}] = \mu_0 [\tilde{\mathbf{H}}_n + \xrightarrow{\rightarrow} \chi_m^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}_n] = \mu_0 [\xrightarrow{\rightarrow} \mathbf{I} + \xrightarrow{\rightarrow} \chi_m^{(1)}(\omega_n)] \cdot \tilde{\mathbf{H}}_n \\
 &= \mu_0 \tilde{\boldsymbol{\mu}}_r^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}_n = \tilde{\boldsymbol{\mu}}^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}_n \text{ ①}
 \end{aligned} \tag{C19}$$

$$\tilde{\mathbf{B}}_n^{(l)} = \mu_0 \tilde{\mathbf{M}}_n^{(l)} \quad (l \geq 2), \quad \tilde{\mathbf{B}}_n^{NL} = \mu_0 \tilde{\mathbf{M}}_n^{NL} = \mu_0 \sum_{l=2}^{\infty} \tilde{\mathbf{M}}_n^{(l)}$$

$$\tilde{\mathbf{B}}_n = \tilde{\mathbf{B}}_n^{(1)} + \tilde{\mathbf{B}}_n^{NL}, \quad \tilde{\mathbf{M}}_n = \tilde{\mathbf{M}}_n^{(1)} + \tilde{\mathbf{M}}_n^{NL} \tag{C20}$$

附C1.2.3.2 引入单色化后的 $\tilde{\mathbf{M}}_n$ 的非线性效应，将其线性部分左移，非线性部分保留
将式(C20)中的 $\tilde{\mathbf{M}}_n = \tilde{\mathbf{M}}_n^{(1)} + \tilde{\mathbf{M}}_n^{NL}$ 及式(C18)中的 $\tilde{\mathbf{M}}_n^{(1)} = \xrightarrow{\rightarrow} \chi_m^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}_n$ 代入(C17)得：

$$\frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n) = \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n^{(1)}) + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n^{NL}) \tag{C21}$$

① 同理， $\tilde{\mathbf{B}}_n = \tilde{\mathbf{B}}_n^{(1)} + \tilde{\mathbf{B}}_n^{NL} = \tilde{\mathbf{D}}_n = \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}}} \xrightarrow{\rightarrow} \tilde{\boldsymbol{\mu}}^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}(\omega_{m_{\pi(1)}}) + \sum_{\pi}^{\omega_n = \omega_{m_{\pi(1)}} + \omega_{m_{\pi(2)}}} D_{m_1, m_2} \xrightarrow{\rightarrow \rightarrow} \tilde{\boldsymbol{\mu}}^{(2)}(\omega_n) \cdot \tilde{\mathbf{H}}(\omega_{m_{\pi(1)}}) \tilde{\mathbf{H}}(\omega_{m_{\pi(2)}}) + \dots$

$$\begin{aligned}
\frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n^{(1)}) &= \frac{\partial}{\partial t} \left(\nabla \times [\vec{\chi}_m^{(1)}(\omega_n) \cdot \tilde{\mathbf{H}}_n] \right) \xrightarrow{\vec{\chi}_m^{(1)}(\omega_n) \text{不含时空}} \vec{\chi}_m^{(1)}(\omega_n) \cdot \frac{\partial}{\partial t} [\nabla \times \tilde{\mathbf{H}}_n] \\
&= \vec{\chi}_m^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{D}}_n}{\partial t^2} = \vec{\chi}_m^{(1)}(\omega_n) \cdot \frac{\partial^2 (\tilde{\mathbf{D}}_n^{(1)} + \tilde{\mathbf{D}}_n^{NL})}{\partial t^2} = \vec{\chi}_m^{(1)}(\omega_n) \cdot \frac{\partial^2 [\vec{\epsilon}^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}_n + \tilde{\mathbf{P}}_n^{NL}]}{\partial t^2} \\
&= \vec{\chi}_m^{(1)}(\omega_n) \cdot [\vec{\epsilon}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{E}}_n}{\partial t^2} + \frac{\partial^2 \tilde{\mathbf{P}}_n^{NL}}{\partial t^2}] \quad (C22)
\end{aligned}$$

于是方程(C17)变成非线性介质中单色电场波动方程的最普遍形式:

$$\begin{aligned}
\nabla \times (\nabla \times \tilde{\mathbf{E}}_n) + \frac{\vec{\mathbf{I}} + \vec{\chi}_m^{(1)}(\omega_n) \cdot \vec{\epsilon}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{E}}_n}{\partial t^2}}{\epsilon_0 c^2} \\
= -\frac{\vec{\mathbf{I}} + \vec{\chi}_m^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{P}}_n^{NL}}{\partial t^2}}{\epsilon_0 c^2} - \frac{1}{\epsilon_0 c^2} \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n^{NL}) \quad (C23)
\end{aligned}$$

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}_n) + \vec{\mu}^{(1)}(\omega_n) \cdot \vec{\epsilon}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{E}}_n}{\partial t^2} = -\vec{\mu}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{P}}_n^{NL}}{\partial t^2} - \mu_0 \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{M}}_n^{NL}) \quad (C24)$$

若忽略介质的非线性磁化 $\tilde{\mathbf{M}}_n^{NL}$ ，只考虑线性磁化项 $\tilde{\mathbf{M}}_n^{(1)}$ 时，上述方程变为:

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}_n) + \vec{\mu}^{(1)}(\omega_n) \cdot \vec{\epsilon}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{E}}_n}{\partial t^2} = -\vec{\mu}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{P}}_n^{NL}}{\partial t^2} \quad (C25)$$

若忽略介质的所有磁化 $\tilde{\mathbf{M}}_n$ ，即介质为非磁介质，为纯电介质时，方程(C17)变为:

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}_n) + \frac{1}{c^2} \vec{\epsilon}_r^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{E}}_n}{\partial t^2} = -\frac{1}{\epsilon_0 c^2} \frac{\partial^2 \tilde{\mathbf{P}}_n^{NL}}{\partial t^2} \quad (C26)$$

附C1.2.3.3 磁场波动方程的单色化和分离线性、非线性部分

同样的操作，可将磁场波动方程(C8)单色化为类似(C17)的:

$$\nabla \times (\nabla \times \tilde{\mathbf{H}}_n) + \frac{1}{c^2} \vec{\mu}_r^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{H}}_n}{\partial t^2} = -\frac{1}{c^2} \frac{\partial^2 \tilde{\mathbf{M}}_n^{NL}}{\partial t^2} + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{P}}_n) \quad (C27)$$

接着可算得:

$$\frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{P}}_n^{(1)}) = -\epsilon_0 \vec{\chi}_e^{(1)}(\omega_n) \cdot [\vec{\mu}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{H}}_n}{\partial t^2} + \mu_0 \frac{\partial^2 \tilde{\mathbf{M}}_n^{NL}}{\partial t^2}] \quad (C28)$$

可得类似式(C24)的非线性介质中单色磁场波动方程的最普遍形式:

$$\nabla \times (\nabla \times \tilde{\mathbf{H}}_n) + \vec{\epsilon}^{(1)}(\omega_n) \cdot \vec{\mu}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{H}}_n}{\partial t^2} = -\mu_0 \vec{\epsilon}^{(1)}(\omega_n) \cdot \frac{\partial^2 \tilde{\mathbf{M}}_n^{NL}}{\partial t^2} + \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{P}}_n^{NL}) \quad (C29)$$

尽管可以继续得到忽略非线性极化 $\tilde{\mathbf{P}}_n^{NL}$ 、只考虑线性极化 $\tilde{\mathbf{P}}_n^{(1)}$ 和忽略所有极化 $\tilde{\mathbf{P}}_n$ 的类似(C25)和(C26)的方程；但在非线性光学中，激光光场在非线性晶体中所极化出的 $\tilde{\mathbf{P}}_n^{NL}$ 的量级可与 $\tilde{\mathbf{H}}_n$ 比拟，所以不能忽略式(C29)中的 $\partial(\nabla \times \tilde{\mathbf{P}}_n^{NL})/\partial t$ ，也不能忽略式

(C28)中的 $\partial(\nabla \times \tilde{\mathbf{P}})/\partial t$ 。

附C1.2.4 单色化之傅里叶级数展开、分离并去掉单色波的时间项、定态波动方程

附C1.2.4.1 数学、物理上的实数、复数形式的傅里叶级数展开

从物理意义上，任何电场矢量 $\tilde{\mathbf{E}}$ （如非单色场、驻波场等）都可写为不连续单色子波的黎曼和；但更深刻地，从数学物理方法的角度，其本质来源于任何一个周期函数都可用基本函数族中的三角函数族展开为傅里叶级数，或用三角函数族的复数形式作为基函数以展开为复数形式的傅里叶级数。比如统计物理中对有限空间中的光子/热辐射系统采取箱归一化或B-K周期性边界条件，就对应着数学上即使是定义在有限区间上的非周期函数 $f(x)$ 也可通过延拓的方法，改造其为另一周期函数 $g(x)$ ，对新周期函数做傅里叶展开后，可用级数和在相应原有限区间上的值，代表原非周期函数；而物理上光波既可以用三角函数族表示，同时也可以用三角函数族的复数形式表示，本质上也就对应着数学上实数形式的傅里叶级数可推导出复数形式的傅里叶级数。

因此据数学上的两种傅里叶展开形式：

$$\begin{aligned} g(x) &= a_0 + \sum_{i=1}^{\infty} \left[a_i \cos\left(\frac{2\pi}{L} i \cdot x\right) + b_i \sin\left(\frac{2\pi}{L} i \cdot x\right) \right] (a, b \in \mathbb{R}) \\ g(x) &= \sum_{i=-\infty}^{\infty} c_i e^{\frac{i 2\pi}{L} i \cdot x} \left(c_0 = a_0, c_+ = \frac{a_i - ib_i}{2}, c_- = \frac{a_i + ib_i}{2}; c \in \mathbb{C} \right) \end{aligned} \quad (\text{C30})$$

相应地，物理上对物理量的傅里叶展开也可分为两种，一种是实数形式的，时域与空域，即下面两式

$$\begin{aligned} \tilde{\mathcal{E}}(t) &= A_0 + \sum_{i=1}^{\infty} \left[A_i \cos\left(\frac{2\pi}{T} i \cdot t\right) + B_i \sin\left(\frac{2\pi}{T} i \cdot t\right) \right] \\ &= A_0 + \sum_{i=1}^{\infty} [A_i \cos(\omega_i \cdot t) + B_i \sin(\omega_i \cdot t)] \end{aligned} \quad (\text{C31})$$

$$\begin{aligned} \tilde{\mathcal{E}}(\mathbf{r}) &= A_0 + \sum_{i=1}^{\infty} \left[A_i \cos\left(\frac{2\pi}{\lambda} i \hat{\mathbf{k}}_i \cdot \mathbf{r}\right) + B_i \sin\left(\frac{2\pi}{\lambda} i \hat{\mathbf{k}}_i \cdot \mathbf{r}\right) \right] \\ &= A_0 + \sum_{i=1}^{\infty} [A_i \cos(\mathbf{k}_i \cdot \mathbf{r}) + B_i \sin(\mathbf{k}_i \cdot \mathbf{r})] \end{aligned} \quad (\text{C32})$$

所合成的：

$$\tilde{\mathcal{E}}(\mathbf{r}, t) = A_0 + \sum_{i=1}^{\infty} [A_i \cos(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t) + B_i \sin(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)]$$

$$(|A|, |B| \in \mathbb{R}, \text{且不要求 } A_i // B_i) \quad (\text{C33})$$

另一种是复数形式的，同时体现点波源时域 $\tilde{\mathcal{E}}(t) = \sum_{i=-\infty}^{\infty} \mathbf{C}_i e^{i\omega_i t}$ 与驻波场空域 $\tilde{\mathcal{E}}(\mathbf{r}) = \sum_{i=-\infty}^{\infty} \mathbf{C}_i e^{i\mathbf{k}_i \cdot \mathbf{r}}$ 的行波场时空区域：

$$\tilde{\mathcal{E}}(\mathbf{r}, t) = \sum_{i=-\infty}^{\infty} \mathbf{C}_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} \left(\mathbf{C}_0 = \mathbf{A}_0, \mathbf{C}_+ = \frac{\mathbf{A}_i - i\mathbf{B}_i}{2}, \mathbf{C}_- = \frac{\mathbf{A}_i + i\mathbf{B}_i}{2}; |\mathbf{C}| \in \mathbb{C} \right) \quad (\text{C34})$$

有的地方也写作：

$$\tilde{\mathcal{E}}(\mathbf{r}, t) = \sum'_i \mathbf{C}_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} + \text{c.c.} = \sum'_i \left(\mathbf{C}_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} + \text{c.c.} \right) \quad (\text{C35})$$

其中， $\omega_0 = 0$ 、 $\mathbf{k}_0 = \mathbf{0}$ ；但对于 $i \neq 0$ ，不要求 $\omega_i = 2\pi/T \cdot i$ 、 $\mathbf{k}_i = 2\pi/\lambda \cdot i\hat{\mathbf{k}}_i$ ^①，只需 ω_i 与 \mathbf{k}_i 的下标 i 都表示某单色光即可， i 不相同的 ω_i 或 \mathbf{k}_i 可以相同。且求和符号上标一撇表示下指标取值遍历所有非零整数；后一种写法中的 c.c. 可视为前一种写法中的直流背景 \mathbf{A}_0 从求和中单独提取出来，并经分解后，分配给前一个写法中每个原不含直流分量的子波 $\mathbf{E}_i = \mathbf{C}_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)}$ ($i \neq 0$)，作为其直流分量成分，合成得后一种写法中的新子波 $\mathbf{E}_i = \mathbf{C}_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} + \text{c.c.}$ ($i \neq 0$)；这种常数分配也可用于实数形式的傅里叶级数。但为方便起见，之后的推导在子波和总波形式上仍采用前一种写法。

附C1.2.4.2 选择合适的坐标系，物理量可展为实数形式的傅里叶余弦级数，且相应复数形式的级数系数为实数

现在希望子波中各振幅 $|\mathbf{C}_i| \in \mathbb{R}$ 而不是 $|\mathbf{C}_i| \in \mathbb{C}$ 。为此，数学上的做法是选择合适的延拓方法，同时选择合适的坐标原点，以使得原函数 $f(x)$ 所延拓成的 $g(x)$ 是个周期函数的同时，在所选择的坐标系和坐标原点下是个偶函数，此时 $g(x)$ 的实的傅里叶级数中的各 $b_i = 0$ ， $g(x)$ 便展成了傅里叶余弦级数，而其复数形式的傅里叶级数中的各 c_i 除了 $c_0 = a_0$ 外，其他 $c_i = a_{|i|}/2$ ($i \neq 0$)，此时所有的 $c \in \mathbb{R}$ 。

类似地，物理上也总能做到选择合适的时空坐标系、时空坐标原点，以及延拓方法，使得延拓后的 $\tilde{\mathcal{E}}(t)$ 、 $\tilde{\mathcal{E}}(\mathbf{r})$ 以及 $\tilde{\mathcal{E}}(\mathbf{r}, t)$ 在四个维度上分别都是偶函数且是周期函数，则此时各 $\mathbf{B}_i = \mathbf{0}$ ，且 $\mathbf{C}_i = \frac{\mathbf{A}_{|i|}}{2}$ ($i \neq 0$)、 $|\mathbf{C}| \in \mathbb{R}$ 。此时

$$\tilde{\mathcal{E}}(\mathbf{r}, t) = \mathbf{A}_0 + \sum_{i=1}^{\infty} \mathbf{A}_i \cos(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t) = \sum_{i=-\infty}^{\infty} \mathbf{C}_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} \quad (\text{C36})$$

^① $\omega_i = \omega_{m_i}$ 可看做集合 $\{\omega_{m_i}\}$ 中的元素； \mathbf{k}_i 与 ω_i 类似，都属于光束的属性、标签之一，所以 \mathbf{k} 也有像 ω 的频谱 $\{\omega_m\}$ 一样的集合：如 $\{\mathbf{k}_s\} \leftrightarrow \{s\} \leftarrow \{i\}$ 、 $\{\mathbf{k}_{s_i}\} \leftrightarrow \{s_i\} \leftarrow \{i\}$ ， $\{\mathbf{k}_s\}, \{s\}, \{\mathbf{k}_{s_i}\}, \{s_i\}$ 四者都可表 $\tilde{\mathbf{E}}$ 的“波矢谱”。

也可写为：

$$\tilde{\mathcal{E}}(\mathbf{r}, t) = \sum_{i=1}^{\infty} [A_i \cos(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t) + \text{c.c.}] = \sum_i' (C_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} + \text{c.c.}) \quad (\text{C37})$$

但不建议这么定义。现将系数改写作：

$$\begin{aligned} \tilde{\mathcal{E}}(\mathbf{r}, t) &= \sum_{i=0}^{\infty} \tilde{\mathcal{E}}_i(\mathbf{r}, t) = \sum_{i=0}^{\infty} \mathcal{E}_i \cos(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t) = \sum_{i=-\infty}^{\infty} A_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} = \sum_{i=-\infty}^{\infty} \tilde{\mathbf{E}}_i(\mathbf{r}, t) \\ &\left(\mathbf{A}_0 = \mathcal{E}_0, \mathbf{A}_i = \frac{\mathcal{E}_i}{2} (i \neq 0); |A| \in \mathbb{R} \right) \end{aligned} \quad (\text{C38})$$

其中每个单色子波（一般只关心复数形式的单色子波） $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 的系数部分和空间部分 $\mathbf{E}_i(\mathbf{r}) = \mathbf{E}(\omega_i, \mathbf{k}_i) = \mathbf{E}(\omega_i)$ ^①，即 $A_i e^{i\mathbf{k}_i \cdot \mathbf{r}} = A(\omega_i, \mathbf{k}_i) e^{i\mathbf{k}_i \cdot \mathbf{r}} = A(\omega_i) e^{i\mathbf{k}_i \cdot \mathbf{r}}$ 或 $\mathcal{E}_i e^{i\mathbf{k}_i \cdot \mathbf{r}} / 2 = \mathcal{E}(\omega_i) e^{i\mathbf{k}_i \cdot \mathbf{r}} / 2$ 部分，为该单色子波 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 的复振幅，而复振幅中的（非相位）振幅部分如 $A_i, A(\omega_i), \mathcal{E}_i / 2, \mathcal{E}(\omega_i) / 2$ ，为单色子波 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 的（实）振幅； $e^{-i\omega_i t}$ 为单色子波 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 的时间部分，只含相位成分，与该 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 的频率 ω_i 有关。

附C1.2.4.3 定义负下标和负频率及对应的物理量，以简化可观测量的数学表示

之所以要将 $\tilde{\mathbf{E}}(\mathbf{r}, t)$ 展成复数形式的级数，除了表示上的便利，一方面便于乘法和乘方运算，另一方面也便于分离参与混频的单色入射波的时间变量；复数形式的单色子波 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 还有如下数学性质：对 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 取复共轭，原本只能进行到这种地步：

$$\tilde{\mathbf{E}}_i(\mathbf{r}, t)^* (= \tilde{\mathbf{E}}(\omega_i, \mathbf{k}_i)^* = \tilde{\mathbf{E}}(\omega_i))^* = \mathbf{E}_i(\mathbf{r})^* e^{-i(-\omega_i)t} \quad (\text{C39})$$

其中 $\mathbf{E}_i(\mathbf{r})^* = \mathbf{E}(\omega_i, \mathbf{k}_i)^* = A(\omega_i, \mathbf{k}_i)^* e^{i(-\mathbf{k}_i) \cdot \mathbf{r}}$ ；由于只定义了非负整数取值的*i*及其相应的 ω_i, \mathbf{k}_i 的物理意义，此处接着定义取值为负整数的*i*及其所代表的波的各项光学参数 ω_i, \mathbf{k}_i 和 A_i, \mathbf{E}_i 连同 $\tilde{\mathbf{E}}_i$ 本身的物理意义：对于*i*>0，定义： $\omega_{-i} := -\omega_i$ 、 $\mathbf{k}_{-i} := -\mathbf{k}_i$ ，于是 $\tilde{\mathbf{E}}_i^* = \tilde{\mathbf{E}}(\omega_i, \mathbf{k}_i)^* = \tilde{\mathbf{E}}_i(\mathbf{r}, t)^* = \mathbf{E}_i(\mathbf{r})^* e^{-i\omega_{-i}t}$ ，再定义

$$A_i^* = A(\omega_i, \mathbf{k}_i)^* := A(-\omega_i, -\mathbf{k}_i) = A(\omega_{-i}, \mathbf{k}_{-i}) = A_{-i} \quad (\text{C40})$$

则有 $\mathbf{E}_i^* = \mathbf{E}(\omega_i, \mathbf{k}_i)^* = A(\omega_i, \mathbf{k}_i)^* e^{i(-\mathbf{k}_i) \cdot \mathbf{r}} = A(\omega_{-i}, \mathbf{k}_{-i}) e^{i\mathbf{k}_{-i} \cdot \mathbf{r}} = \mathbf{E}(\omega_{-i}, \mathbf{k}_{-i}) = \mathbf{E}_{-i}$ ，反代回 $\tilde{\mathbf{E}}_i^*$ 得

^① 根据 $\tilde{\mathbf{E}}_i$ 的表达式 $A_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)}$ ，最推荐的两种写法应为 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 和 $\tilde{\mathbf{E}}(\omega_i, \mathbf{k}_i)$ ，写 $\tilde{\mathbf{E}}(\omega_i)$ 以及不写 $\tilde{\mathbf{E}}(\mathbf{k}_i)$ 完全是由于重 ω_i 轻 \mathbf{k}_i 的思想，本应不用这两种写法；同理，根据 \mathbf{E}_i 的表达式 $A_i e^{i\mathbf{k}_i \cdot \mathbf{r}}$ ，最推荐的两种写法为 $\mathbf{E}(\mathbf{k}_i)$ 和 $\mathbf{E}_i(\mathbf{r})$ ，且本不应用 $\mathbf{E}(\omega_i, \mathbf{k}_i)$ 的写法。但人们总关心频率这一特征，就用 $\mathbf{E}(\omega_i)$ 替换掉了 $\mathbf{E}(\mathbf{k}_i)$ ；同理，若认为 A_i 属于 $\tilde{\mathbf{E}}_i$ ，则可以用 $A(\omega_i, \mathbf{k}_i)$ 这种写法。而若认为 A_i 属于 \mathbf{E}_i ，则采用 $A(\mathbf{k}_i)$ ，但人们又总关心光束的频率，所以又用 $A(\omega_i)$ 替换掉了 $A(\mathbf{k}_i)$ 。这里要注意无论如何 $A_i(\cdot)$ 中不应含 \mathbf{r} 或 t 。

$$\tilde{\mathbf{E}}_i^* = \tilde{\mathbf{E}}(\omega_i, \mathbf{k}_i)^* = \mathbf{E}(\omega_{-i}, \mathbf{k}_{-i}) e^{-i\omega_{-i}t} = \tilde{\mathbf{E}}(\omega_{-i}, \mathbf{k}_{-i}) = \tilde{\mathbf{E}}_{-i}.$$

这样定义的目的在于原具有物理意义的实数形式单色子波本只能写为

$$\tilde{\mathcal{E}}_i(\mathbf{r}, t) = \mathcal{E}_i \cos(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t) = A_i e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} + A_i e^{-i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)} (i \neq 0), \text{ 但现在可以直接写作}$$

$$\tilde{\mathcal{E}}_i = \tilde{\mathbf{E}}_i + \tilde{\mathbf{E}}_{-i} = \tilde{\mathbf{E}}_i + \tilde{\mathbf{E}}_i^* (i \neq 0) \quad \mathcal{E}_0 = \tilde{\mathbf{E}}_0 \quad (\text{C41})$$

为之后单色能流密度等可观测量的写法加以简化埋下伏笔。

附C1.2.4.4 非线性过程合成的极化光的单色各阶非线性效应展开和复色傅里叶展开

(1) 同频单色入射光集 $\{\tilde{\mathbf{E}}_i|_{\omega_i=\omega}\}$ 所参与的多个 l 阶倍频过程所生成的复色极化光的非线性阶次展开同时也是其傅里叶级数展开

现考虑“单 $\tilde{\mathbf{E}}_i \rightarrow$ 多 $\tilde{\mathbf{P}}_n$ ”过程，对于任意某个频率为 ω_i 的单色混频成分 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ ，其在某一相同的 \mathbf{r} 处，可能引发介质的多倍频非线性响应，产生单色倍频极化光集 $\{\tilde{\mathbf{P}}_n(\mathbf{r}, t)\}$ ，相应频率集为 $\{\omega_n\} = \{l\omega\} = \{n\omega\}$ ，其中可认为 $n=l$ 由 l 控制；这意味着每个通过式(C1)的第 l 阶非线性过程合成的 $\tilde{\mathbf{P}}_n(\mathbf{r}, t)$ 的频率都为 $l=n$ 个 ω 之和，即此时

$\tilde{\mathbf{P}}_n(\mathbf{r}, t) = \tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t) = \mathbf{P}_n^{(l)}(\mathbf{r}) e^{-i\omega_n t} = \mathbf{P}_n^{(n)}(\mathbf{r}) e^{-i\omega_n t} = \tilde{\mathbf{P}}_n^{(n)}(\mathbf{r}, t)$ ^① 的时间项为 $e^{-i\omega_n t} = e^{-il\omega t} = (e^{-i\omega t})^l = (e^{-i\omega t})^n = e^{-in\omega t}$ ，这种情况下，由不同的 l 所对应的，通过第 l 阶非线性过程合成的 $\tilde{\mathbf{P}}_n(\mathbf{r}, t)$ ，再合成的极化波叠加场：

$$\tilde{\mathbf{P}}(\mathbf{r}, t) = \sum_n \tilde{\mathbf{P}}_n(\mathbf{r}, t) = \sum_{l=1}^{\infty} \tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t) = \sum_{n=1}^{\infty} \tilde{\mathbf{P}}_n^{(n)}(\mathbf{r}, t) \quad (\text{C42})$$

(2) 复色入射光集 $\{\tilde{\mathbf{E}}_i\}$ 通过多个 l 阶和/差频过程生成的复色极化光和电位移矢量的两种展开

再考虑“多 $\tilde{\mathbf{E}}_i \rightarrow$ 单 $\tilde{\mathbf{P}}_n$ ”过程，考虑某一频率为 ω_n 的单色和频极化光，它作为目标即混频/合成的结果，本身就是由式(C1)中各阶非线性效应 $\tilde{\mathbf{P}}_n^{(l)}$ 叠加而成，即

$\tilde{\mathbf{P}}_n(\mathbf{r}, t) = \sum_{l=1}^{\infty} \tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t)$ ^②，而对于其中每一个 $\tilde{\mathbf{P}}_n^{(l)}$ 成分，都有 l 束满足能量守恒 $\sum_{i=1}^l \omega_i = \omega_n$ 的单色 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 参与合成，因此每个合成的极化光 $\tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t) = \mathbf{P}_n^{(l)}(\mathbf{r}) e^{-i\omega_n t}$ 的时间部分为 $e^{-i\omega_n t} = e^{-i\sum_{i=1}^l \omega_i t} = \prod_{i=1}^l e^{-i\omega_i t}$ ，其连同频率 ω_n 均不随阶数 l 的变化而变化，是个固定值。因此合成的 $\tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t)$ 再合成所得的叠加场（虽是叠加场，但仍是单色的）

^① 也可写为诸如 $\tilde{\mathbf{P}}_l(\mathbf{r}, t) = \tilde{\mathbf{P}}_l^{(l)}(\mathbf{r}, t) = \mathbf{P}_l^{(l)}(\mathbf{r}) e^{-i\omega_n t}$ 等，但表意不明所以不推荐：因为下标应是 n 代入了 $n=l$ 后才变为 l 的；注意其中 $\tilde{\mathbf{P}}_n^{(l)}$ 的上标值会跟着下标值的变动而变动，下标值也会跟着上标值变动而变动。

^② 该求和中的 $\tilde{\mathbf{P}}_n^{(l)}$ 与“单 $\tilde{\mathbf{E}}_i \rightarrow$ 多 $\tilde{\mathbf{P}}_n$ ”过程中的相同求和中的 $\tilde{\mathbf{P}}_n^{(l)}$ 不同，其下标 n 不随 l 的值的变动而变动。

$\tilde{\mathbf{P}}_n(\mathbf{r}, t) = \sum_{l=1}^{\infty} \tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t) = e^{-i\omega_n t} \sum_{l=1}^{\infty} \mathbf{P}_n^{(l)}(\mathbf{r}) = \mathbf{P}_n(\mathbf{r}) e^{-i\omega_n t}$ 的时间部分，是所有单色 $\tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t)$ 的公因式 $e^{-i\omega_n t}$ ，也是随所考虑的频率即 n 的选定而固定了的。

接着考虑多个“多 $\tilde{\mathbf{E}}_i \rightarrow$ 单 $\tilde{\mathbf{P}}_n$ ”过程，合成多个不同频率 ω_n （等价于不同光束 ω_{n_j} ）的和/差频极化光 $\tilde{\mathbf{P}}_n(\mathbf{r}, t)$ ，它所叠加起来的复合场便与式(C38)中的 $\tilde{\mathcal{E}}(\mathbf{r}, t) = \sum_{i=-\infty}^{\infty} \tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 拥有同等地位：

$$\begin{aligned} \tilde{\mathcal{P}}(\mathbf{r}, t) &= \sum_{n=-\infty}^{\infty} \tilde{\mathbf{P}}_n(\mathbf{r}, t) = \sum_{n=-\infty}^{\infty} \sum_{l=1}^{\infty} \tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t) \\ &= \sum_{n=-\infty}^{\infty} e^{-i\sum_{i=1}^l \omega_i t} \sum_{l=1}^{\infty} \mathbf{P}_n^{(l)}(\mathbf{r}) = \sum_{n=-\infty}^{\infty} \mathbf{P}_n(\mathbf{r}) e^{-i\omega_n t} \end{aligned} \quad (\text{C43})$$

同理，据各阶单色非线性电位移矢量与相应阶的非线性单色极化波的关系：

$$\tilde{\mathbf{D}}_n^{(l)} = \tilde{\mathbf{P}}_n^{(l)} (l \geq 2) \quad \tilde{\mathbf{D}}_n^{(1)} = \overset{\leftrightarrow}{\mathbf{\epsilon}}^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}_n, \quad \text{有：}$$

$$\begin{aligned} \tilde{\mathcal{D}}(\mathbf{r}, t) &= \sum_{n=-\infty}^{\infty} \tilde{\mathbf{D}}_n(\mathbf{r}, t) = \sum_{n=-\infty}^{\infty} \sum_{l=1}^{\infty} \tilde{\mathbf{D}}_n^{(l)}(\mathbf{r}, t) \\ &= \sum_{n=-\infty}^{\infty} e^{-i\sum_{i=1}^l \omega_i t} \sum_{l=1}^{\infty} \mathbf{D}_n^{(l)}(\mathbf{r}) = \sum_{n=-\infty}^{\infty} \mathbf{D}_n(\mathbf{r}) e^{-i\omega_n t} \end{aligned} \quad (\text{C44})$$

到此为止，可以说已经做到了两大 $\tilde{\mathbf{E}}$ 、 $\tilde{\mathbf{P}}$ 、 $\tilde{\mathbf{D}}$ 阵营中，复合场 $\tilde{\mathcal{E}}(\mathbf{r}, t)$ 、 $\tilde{\mathcal{P}}(\mathbf{r}, t)$ 、 $\tilde{\mathcal{D}}(\mathbf{r}, t)$ 的一级分解，为之前的单色化做出了理论解释；也给出了至最底层级的 $\tilde{\mathcal{P}}(\mathbf{r}, t)$ 、 $\tilde{\mathcal{D}}(\mathbf{r}, t)$ 的末梢式分解，辨明了其组成成分；同时分离了单色化的 $\tilde{\mathbf{E}}_i(\mathbf{r}, t) = \mathbf{E}_i(\mathbf{r}) e^{-i\omega_i t}$ 、单色极化光 $\tilde{\mathbf{P}}_n(\mathbf{r}, t) = \mathbf{P}_n(\mathbf{r}) e^{-i\omega_n t}$ 和单色电位移矢量 $\tilde{\mathbf{D}}_n(\mathbf{r}, t) = \mathbf{D}_n(\mathbf{r}) e^{-i\omega_n t}$ 的时间部分。

(3) 复数非线性系数造成 $\tilde{\mathbf{E}}_i$ 的振幅、能量衰减，并造成 $\tilde{\mathbf{P}}_n$ 、 $\tilde{\mathbf{D}}_n$ 相位延迟及损耗

只不过当入射波频率与介质的固有频率相近时，或者说介质有损耗、对辐射场有吸收时，各阶非线性系数 $\overset{\leftrightarrow}{\chi}_e^{(l)}(\omega_n)$ 、 $\overset{\leftrightarrow}{\epsilon}_r^{(l)}(\omega_n)$ 、 $\overset{\leftrightarrow}{\epsilon}^{(l)}(\omega_n)$ 都是复数张量，其各个分量都不一定是实数了，导致每个参与混频的 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 的各个 i, j, k 三分量，在与三大张量系数矩阵相乘以参与混频的过程中，会先乘以复数即发生相位延迟，再加和起来即发生混频。这对 $\tilde{\mathbf{P}}_n(\mathbf{r}, t)$ 、 $\tilde{\mathbf{D}}_n(\mathbf{r}, t)$ 的直接影响就是其复振幅部分/空间部分即 $\mathbf{P}_n(\mathbf{r})$ 、 $\mathbf{D}_n(\mathbf{r})$ ，其中的相位部分，会引入额外的不含时也不含坐标量的相位因子（但最好还是将此相位因子/复数放入实振幅部分中，使其变为复振幅中的复振幅，或者单独作为一个因子，以保持复振幅中相位部分中的波矢部分的独立性），效果是 $\tilde{\mathbf{P}}_n(\mathbf{r}, t)$ 、 $\tilde{\mathbf{D}}_n(\mathbf{r}, t)$ 整体会多乘上一个模大于 1（对应在原实部基础上添加虚部）的复数，以表示介质产生的响应与介质接受的输入不同步，相位会滞后（对应非线性系数的虚部为正）于参与混频的各电场 $\tilde{\mathbf{E}}_i(\mathbf{r}, t)$ 的相位，而时间部分保持不变；非线性系数为复数还有另一个重要的影响，以各向同性的线性介质之金属(导体)为例，作为典型的耗散介质，其麦氏方程组的四

号方程右侧传导电流 $\tilde{\mathbf{J}}_f$ 不可忽略，导致除位移电流所产生的非线性系数的实部之外，还产生非线性系数的虚部 $i\sigma/\omega$ ，为保持波动方程形式不变，则它将导致波动方程的解中波矢也出现虚部，以至最终解 $\tilde{\mathbf{E}}_n$ 中的复波矢的虚部乘以虚单位 i 后转变为随空间距离衰减的实振幅的负指数，决定了电磁波的衰减特性，可理解为欧姆定律 $\tilde{\mathbf{J}}_f = \sigma \cdot \tilde{\mathbf{E}}$ 外加传导电流的热效应导致 $\tilde{\mathbf{E}}_n$ 偏振方向上能量损耗的发生。因此在导体中，由 $\tilde{\mathbf{E}}_n$ 所极化出的 $\tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 不仅相位会延迟，振幅也会衰减。

(4) $\{\tilde{\mathbf{E}}_i\}$ 到 $\tilde{\mathbf{P}}_n^{NL}, \tilde{\mathbf{D}}_n^{NL}$ 非线性极化过程均满足能量守恒和动量守恒

另外，既然 $\tilde{\mathbf{P}}_n^{(l)}$ 与 $\tilde{\mathbf{D}}_n^{(l)} = \tilde{\mathbf{P}}_n^{(l)} (l \geq 2)$ 中的时间部分是通过式(C1)以 $e^{-i\omega_n t} = e^{-i\sum_{i=1}^l \omega_i t} = \prod_{i=1}^l e^{-i\omega_i t}$ 的方式产生的，而其空间部分中的相位部分，即波矢部分，由于与时间部分即频率部分的数学形式相同、地位相同（比如若写成三角函数形式则它们都是自变量中的相位部分），那么在混频过程中，若同一束 l 阶单色极化波的频率部分由 $e^{-i\omega_n t} = e^{-i\sum_{i=1}^l \omega_i t} = \prod_{i=1}^l e^{-i\omega_i t}$ 给定，则其波矢部分也将由光波序列同为 $\{i\}$ 的相同的 l 束单色光的波矢部分相乘而得： $e^{-ik_n \cdot r} = e^{-i\sum_{i=1}^l k_i \cdot r} = \prod_{i=1}^l e^{-ik_i \cdot r}$ ，因此在 $\tilde{\mathbf{P}}_n^{(l)}$ 与 $\tilde{\mathbf{D}}_n^{(l)}$ 的生成过程中，这 $l+1$ 束波除了满足能量/频率守恒 $\sum_{i=1}^l \omega_i = \omega_n$ 外，还满足动量/波矢守恒 $\sum_{i=1}^l \mathbf{k}_i = \mathbf{k}_n$ ，也就是说 $\tilde{\mathbf{P}}_n^{(l)}$ 与 $\tilde{\mathbf{D}}_n^{(l)}$ 的复振幅部分 $\mathbf{P}_n^{(l)}(\mathbf{r}), \mathbf{D}_n^{(l)}(\mathbf{r})$ 中必然含有相位因子 $e^{-ik_n \cdot r}$ ，且其中波矢 $\mathbf{k}_n = \sum_{i=1}^l \mathbf{k}_i$ ；同理， $\tilde{\mathbf{P}}_n^{NL}, \tilde{\mathbf{D}}_n^{NL}$ 及 $\mathbf{P}_n^{NL}(\mathbf{r}), \mathbf{D}_n^{NL}(\mathbf{r})$ 其中也应均含有 $e^{-ik_n \cdot r}$ 且 $\mathbf{k}_n = \sum_{i=1}^l \mathbf{k}_i$ ；按理说 $\tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 及 $\mathbf{P}_n(\mathbf{r}), \mathbf{D}_n(\mathbf{r})$ 也应是相同的情况，但由于其中包含 $\tilde{\mathbf{P}}_n^{(1)}, \tilde{\mathbf{D}}_n^{(1)}$ ，而它们中又含有 $\tilde{\mathbf{E}}_n$ ，所以对于 $\tilde{\mathbf{E}}_n, \tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ ，事情要稍微复杂一点。

(5) $\tilde{\mathbf{E}}_n$ 线性极化到 $\tilde{\mathbf{P}}_n^{(1)}, \tilde{\mathbf{D}}_n^{(1)}$ 也满足能量动量守恒，但三者不一定满足完美波矢匹配

当 $\tilde{\mathbf{E}}_n$ 中的时间部分满足 $e^{-i\omega_n t} = e^{-i\sum_{i=1}^l \omega_i t}$ ，且空间部分满足 $e^{-ik_n \cdot r} = e^{-i\sum_{i=1}^l k_i \cdot r}$ 时，称为(符合)完美相位匹配条件，但一般情况下只可能符合其中一个条件；为将波动方程化为定态以便求解，一般先假设非线性波源所生成的和频波的频率满足 $\omega_n = \sum_{i=1}^l \omega_i$ ，并于波动方程化为定态且解出 $\tilde{\mathbf{E}}_n$ 后，再考察波矢匹配与否以及相应的结果，即是否有 $\tilde{\mathbf{E}}_n$ 中的 $e^{-ik_n \cdot r}$ 中的 $\mathbf{k}_n = \sum_{i=1}^l \mathbf{k}_i$ ，此时其 \mathbf{k}_n 却不一定 $= \sum_{i=1}^l \mathbf{k}_i$ 。之后我们会给出，这是一个条件，而不是一个结论或事实；相反，要同时实现波矢和频率匹配，即完美相位匹配，是极其苛刻的。只有在完美相位匹配条件下，和频波 $\tilde{\mathbf{E}}_n$ 的振幅才会随着相应非线性和频极化光在非线性介质中的传播而持续线性增大，且同时 $\tilde{\mathbf{E}}_n$ 的波矢才精确地为 $\sum_{i=1}^l \mathbf{k}_i$ 。

在求和 $\tilde{\mathbf{P}} = \sum_{n=-\infty}^{\infty} \tilde{\mathbf{P}}_n$ 、 $\tilde{\mathbf{D}} = \sum_{n=-\infty}^{\infty} \tilde{\mathbf{D}}_n$ 中，也定义了负角标所对应的负频率、负波矢以及振幅部分、复振幅部分和相应时变物理量的共轭量：

$$\tilde{\mathbf{P}}(\omega_n, \mathbf{k}_n)^* = \mathbf{P}(-\omega_n, -\mathbf{k}_n) e^{-i(-\omega_n)t} = \tilde{\mathbf{P}}(\omega_{-n}, \mathbf{k}_{-n}) \quad (\text{C45})$$

$$\tilde{\mathbf{D}}(\omega_n, \mathbf{k}_n)^* = \mathbf{D}(-\omega_n, -\mathbf{k}_n) e^{-i(-\omega_n)t} = \tilde{\mathbf{D}}(\omega_{-n}, \mathbf{k}_{-n}) \quad (\text{C46})$$

包括为了将方程变为定态而将方程的解的频率设为与 $\tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 的频率 ω_n 相同的 $\tilde{\mathbf{E}}_n = \mathbf{E}_n(r) e^{-i\omega_n t}$ ，也定义有：

$$\tilde{\mathbf{E}}(\omega_n, \mathbf{k}_n)^* = \mathbf{E}(-\omega_n, -\mathbf{k}_n) e^{-i(-\omega_n)t} = \tilde{\mathbf{E}}(\omega_{-n}, \mathbf{k}_{-n}) \quad (\text{C47})$$

要注意 $\tilde{\mathbf{E}}_n$ 中的 \mathbf{k}_n 与 $\tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 和 $\tilde{\mathbf{P}}_n^{NL}, \tilde{\mathbf{D}}_n^{NL}$ 中的 \mathbf{k}_n 含义不同，且值可能相同或不同。

(6) 类似 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 的负频和线性叠加，有 $\tilde{\mathbf{H}}_i, \tilde{\mathbf{M}}_n, \tilde{\mathbf{B}}_n$ 的负频和线性叠加，且电磁的非线性过程是同时发生的

同理，对于 $\tilde{\mathbf{H}}, \tilde{\mathbf{M}}, \tilde{\mathbf{B}}$ ，也有 $\tilde{\mathbf{H}}^* = \tilde{\mathbf{H}}_{-n}$ 、 $\tilde{\mathbf{M}}^* = \tilde{\mathbf{M}}_{-n}$ 、 $\tilde{\mathbf{B}}^* = \tilde{\mathbf{B}}_{-n}$ ，以及三个叠加式：

$$\tilde{\mathcal{H}}(\mathbf{r}, t) = \sum_{i=-\infty}^{\infty} \tilde{\mathbf{H}}_i(\mathbf{r}, t) \quad (\text{C48})$$

$$\begin{aligned} \tilde{\mathcal{M}}(\mathbf{r}, t) &= \sum_{n=-\infty}^{\infty} \tilde{\mathbf{M}}_n(\mathbf{r}, t) = \sum_{n=-\infty}^{\infty} \sum_{l=1}^{\infty} \tilde{\mathbf{M}}_n^{(l)}(\mathbf{r}, t) \\ &= \sum_{n=-\infty}^{\infty} e^{-i \sum_{i=1}^l \omega_i t} \sum_{l=1}^{\infty} \mathbf{M}_n^{(l)}(\mathbf{r}) = \sum_{n=-\infty}^{\infty} \mathbf{M}_n(\mathbf{r}) e^{-i \omega_n t} \end{aligned} \quad (\text{C49})$$

$$\begin{aligned} \tilde{\mathcal{B}}(\mathbf{r}, t) &= \sum_{n=-\infty}^{\infty} \tilde{\mathbf{B}}_n(\mathbf{r}, t) = \sum_{n=-\infty}^{\infty} \sum_{l=1}^{\infty} \tilde{\mathbf{B}}_n^{(l)}(\mathbf{r}, t) \\ &= \sum_{n=-\infty}^{\infty} e^{-i \sum_{i=1}^l \omega_i t} \sum_{l=1}^{\infty} \mathbf{B}_n^{(l)}(\mathbf{r}) = \sum_{n=-\infty}^{\infty} \mathbf{B}_n(\mathbf{r}) e^{-i \omega_n t} \end{aligned} \quad (\text{C50})$$

其中由于式(C1)与式(C18)的每一阶非线性效应均共享同一组参与混频的电磁场序列 $(\pi(1), \pi(2), \dots, \pi(l))$ ，以至 $\tilde{\mathbf{M}}_n^{(l)}$ 与 $\tilde{\mathbf{B}}_n^{(l)} = \mu_0 \tilde{\mathbf{M}}_n^{(l)} (l \geq 2)$ 在合成过程中使用的原材料，是与 $\tilde{\mathbf{P}}_n^{(l)}$ 与 $\tilde{\mathbf{D}}_n^{(l)} = \tilde{\mathbf{P}}_n^{(l)} (l \geq 2)$ 所使用的原材料相同的 l 束波，的磁场部分 $(\tilde{\mathbf{H}}_1, \tilde{\mathbf{H}}_2, \dots, \tilde{\mathbf{H}}_l)$ 和电场部分 $(\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2, \dots, \tilde{\mathbf{E}}_l)$ ^①；因此不仅式(C1)与式(C18)的相同阶的非线性过程同时发生，且共享同一能量守恒过程 $\sum_{i=1}^l \omega_i = \omega_n$ 和动量守恒过程 $\sum_{i=1}^l \mathbf{k}_i = \mathbf{k}_n$ ；不仅同时合成出

^① 每一束波的 $\tilde{\mathbf{E}}_i$ 与 $\tilde{\mathbf{H}}_i$ 的波矢和频率相同，由四个麦氏方程程导出的电场、磁场两个波动方程的两个解保证。

$\tilde{\mathbf{M}}_n^{(l)}$ 与 $\tilde{\mathbf{P}}_n^{(l)}$ ，且二者的频率、波矢均相同^①；然后无数阶这样的非线性过程的结果之和 $\tilde{\mathbf{M}}_n^{NL}$ 与 $\tilde{\mathbf{P}}_n^{NL}$ （频率、波矢也均相同），再作为各自的波动方程右侧的非线性波源项，产生各自对应的和/差频波 $\tilde{\mathbf{H}}_n$ 与 $\tilde{\mathbf{E}}_n$ ，二者的频率共享 ω_n 且值仍为 $\sum_{i=1}^l \omega_i$ ，波矢也共享同一个 \mathbf{k}_n ，但值不一定为 $\sum_{i=1}^l \mathbf{k}_i$ ；接着 $\tilde{\mathbf{H}}_n$ 与 $\tilde{\mathbf{E}}_n$ 再通过一阶(非)线性效应，合成出 $\tilde{\mathbf{M}}_n^{(1)}$ 与 $\tilde{\mathbf{P}}_n^{(1)}$ 以及 $\tilde{\mathbf{B}}_n^{(1)} = \mu_0[\tilde{\mathbf{H}}_n + \tilde{\mathbf{M}}_n^{(1)}]$ 与 $\tilde{\mathbf{D}}_n^{(1)} = \epsilon_0 \tilde{\mathbf{E}}_n + \tilde{\mathbf{P}}_n^{(1)}$ ；最后合成出麦氏方程组中与 $\tilde{\mathbf{H}}_n$ 与 $\tilde{\mathbf{E}}_n$ 等价的 $\tilde{\mathbf{M}}_n = \tilde{\mathbf{M}}_n^{(1)} + \tilde{\mathbf{M}}_n^{NL}$ ， $\tilde{\mathbf{P}}_n = \tilde{\mathbf{P}}_n^{(1)} + \tilde{\mathbf{P}}_n^{NL}$ 、 $\tilde{\mathbf{B}}_n = \tilde{\mathbf{B}}_n^{(1)} + \tilde{\mathbf{B}}_n^{NL}$ ， $\tilde{\mathbf{D}}_n = \tilde{\mathbf{D}}_n^{(1)} + \tilde{\mathbf{D}}_n^{NL}$ ，至此麦氏方程组中六个物理量的所有阶次的值以及得来顺序都已大白。当然，有磁损耗的介质中的 $\tilde{\boldsymbol{\chi}}_m^{(l)}(\omega_n)$ ， $\tilde{\boldsymbol{\mu}}_r^{(l)}(\omega_n)$ ， $\tilde{\boldsymbol{\mu}}^{(l)}(\omega_n)$ 也都是复数张量。

(7) $\tilde{\mathbf{E}}_n, \tilde{\mathbf{H}}_n, \tilde{\mathbf{P}}_n^{NL}, \tilde{\mathbf{M}}_n^{NL}$ 均单色化并分离时间项后，电磁时变波动方程变为定态

据以上理论，方程(C24)中各三维矢量均可分离出相同的时间部分 $e^{-i\omega_n t}$ ：

$$\begin{aligned}\tilde{\mathbf{E}}_n &= \tilde{\mathbf{E}}_n(\mathbf{r}, t) = \mathbf{E}_n(\mathbf{r})e^{-i\omega_n t} \text{ ②} , \quad \tilde{\mathbf{P}}_n^{NL} = \sum_{l=2}^{\infty} \tilde{\mathbf{P}}_n^{(l)}(\mathbf{r}, t) = e^{-i\sum_{i=1}^l \omega_i t} \sum_{l=2}^{\infty} \mathbf{P}_n^{(l)}(\mathbf{r}) = \mathbf{P}_n^{NL}(\mathbf{r})e^{-i\omega_n t} , \\ \tilde{\mathbf{M}}_n^{NL} &= \mathbf{M}_n^{NL}(\mathbf{r})e^{-i\omega_n t} , \text{ 所以时变方程(C24)、(C25)、(C26)两边均可去掉时间项，变为定} \\ \text{态波动方程，即有下面的：}\end{aligned}$$

$$\begin{aligned}\nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})] - \omega_n^2 \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)}(\omega_n) \cdot \overset{\rightarrow}{\boldsymbol{\varepsilon}}^{(1)}(\omega_n) \cdot \mathbf{E}_n(\mathbf{r}) \\ = \omega_n^2 \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)}(\omega_n) \cdot \mathbf{P}_n^{NL}(\mathbf{r}) + \omega_n^2 \mu_0 [\nabla \times \mathbf{M}_n^{NL}(\mathbf{r})] \quad (C51)\end{aligned}$$

$$\nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})] - \omega_n^2 \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)}(\omega_n) \cdot \overset{\rightarrow}{\boldsymbol{\varepsilon}}^{(1)}(\omega_n) \cdot \mathbf{E}_n(\mathbf{r}) = \omega_n^2 \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)}(\omega_n) \cdot \mathbf{P}_n^{NL}(\mathbf{r}) \quad (C52)$$

$$\nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})] - \frac{\omega_n^2}{c^2} \overset{\rightarrow}{\boldsymbol{\varepsilon}}_r^{(1)}(\omega_n) \cdot \mathbf{E}_n(\mathbf{r}) = \frac{\omega_n^2}{\epsilon_0 c^2} \mathbf{P}_n^{NL}(\mathbf{r}) \quad (C53)$$

式(C53)即三维空间中无磁非线性电介质中的定态波动方程；同理有磁场方面的：

$$\begin{aligned}\nabla \times [\nabla \times \mathbf{H}_n(\mathbf{r})] - \omega_n^2 \overset{\rightarrow}{\boldsymbol{\varepsilon}}^{(1)}(\omega_n) \cdot \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)}(\omega_n) \cdot \mathbf{H}_n(\mathbf{r}) \\ = \omega_n^2 \mu_0 \overset{\rightarrow}{\boldsymbol{\varepsilon}}^{(1)}(\omega_n) \cdot \mathbf{M}_n^{NL}(\mathbf{r}) - \omega_n^2 [\nabla \times \mathbf{P}_n^{NL}(\mathbf{r})] \quad (C54)\end{aligned}$$

附C1.2.5 各向异性的电介质中的定态波动方程及其电矢量无源近似解

^① 既由电磁的非线性过程原料角标和各角标的 $\tilde{\mathbf{E}}_i$ 与 $\tilde{\mathbf{H}}_i$ 的波矢和频率相同，也由麦氏方程组或波动方程约束。

^② 在式(C15)之前提到过，和频波 $\tilde{\mathbf{E}}_n$ 可看做无限个 l 阶非线性效应中， $\{\tilde{\mathbf{E}}_i\}$ 之外的第 $l+1$ 个参与混频的合成波 $\tilde{\mathbf{E}}_{l+1}$ 的加和，而 $\tilde{\mathbf{E}}_{l+1}$ 属于 $\tilde{\mathbf{E}}_i|_{i=l+1}$ ，因此单色 $\tilde{\mathbf{E}}_i(\mathbf{r}, t) = \mathbf{E}_i(\mathbf{r})e^{-i\omega_i t}$ 时间项分离，可应用于单色 $\tilde{\mathbf{E}}_n$ 。

附C1.2.5.1 $\overset{\rightarrow}{\chi}_e^{(1)}(\omega_n), \overset{\rightarrow}{\chi}_m^{(1)}(\omega_n)$ 均对角化后，各向同性、各向异性非线性电介质中的波动方程

各向异性、各向同性一般均考察的是介质的二阶极化率张量（一阶非线性系数） $\overset{\rightarrow}{\chi}_e^{(1)}(\omega_n)$ 。一般的各向异性电介质，比如对称性比较低的晶体，在一般的坐标系下， $\overset{\rightarrow}{\chi}_e^{(1)}(\omega_n)$ 的非主对角元素不全为零，若电介质同时也是非线性的，则其独立分量数为 9，但在线性介质中退化为对称张量（参见附 D），独立分量数为 6；在特殊的坐标系下，即使是一般的各向异性电介质的一般的 $\overset{\rightarrow}{\chi}_e^{(1)}(\omega_n)$ ，不论介质是否同时是非线性的，只要其有三个线性无关的特征向量，均可通过坐标轴变换所驱动的坐标变换而主轴化或对角化，以至于总可将其非主对角元素全置零，独立分量数目变为 3。但接下来的所有物理量，包括麦氏方程组中描述电磁波属性的六个三维矢量，以及二阶磁化率张量 $E_n(r) - [\hat{k}_n \cdot E_n(r)] \hat{k}_n = \frac{1}{\varepsilon_n^{(1)}} \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot E_n(r)$ ，均需要像 $\overset{\rightarrow}{\chi}_e^{(1)}(\omega_n)$ 一样，参照此坐标系。

现暂无证据和理论表明当 $\overset{\rightarrow}{\chi}_e^{(1)}(\omega_n)$ 对角化后， $\overset{\rightarrow}{\chi}_m^{(1)}(\omega_n)$ 也同时一并对角化，即极化率椭球的三个主轴不一定与磁化率椭球的三个主轴重合，选择合适的坐标系只能主轴化二者中的一个。但如此一来即使选取合适的坐标系且 $\nabla \cdot E_n(r) = 0$ ，也无法在电磁均各向异性的电磁介质中，使式(C52)的分量方程左侧只含 $E_n(r)$ 在该方向的分量；只有当介质是电各向同性、磁各向异性，或磁各向同性、电各向异性，或电磁均各向同性时，分量方程才可能不含其它方向的分量，并解出平面电磁波。

现假设二者可同时被对角化，即两个二阶张量的三个主轴两两重合（但长短轴不必一一对应）；当选择了合适的坐标系以将 $\overset{\rightarrow}{\chi}_e^{(1)}(\omega_n), \overset{\rightarrow}{\chi}_m^{(1)}(\omega_n)$ 均对角化后，式(C52)可写为由坐标系的三个坐标轴方向上的三个分量方程，所构成的方程组，并写出其中一个分量方程的具体形式：

$$\begin{aligned} & \left| \begin{array}{ccc} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial z} & \end{array} \right| \left| \begin{array}{c} \mu_{nxx}^{(1)} \epsilon_{nxx}^{(1)} \mathbf{ii} \\ \mu_{nyy}^{(1)} \epsilon_{nyy}^{(1)} \mathbf{jj} \\ \mu_{nzz}^{(1)} \epsilon_{nzz}^{(1)} \mathbf{kk} \end{array} \right| \cdot \begin{bmatrix} E_{nx}(r) \\ E_{ny}(r) \\ E_{nz}(r) \end{bmatrix} \\ & = \omega_n^2 \begin{bmatrix} \mu_{nxx}^{(1)} P_{nx}^{NL}(r) \\ \mu_{nyy}^{(1)} P_{ny}^{NL}(r) \\ \mu_{nzz}^{(1)} P_{nz}^{NL}(r) \end{bmatrix} \quad (C55) \end{aligned}$$

$$\begin{aligned} & \frac{\partial}{\partial y} \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ E_{nx}(\mathbf{r}) & E_{ny}(\mathbf{r}) \end{vmatrix} - \frac{\partial}{\partial z} \begin{vmatrix} \frac{\partial}{\partial z} & \frac{\partial}{\partial x} \\ E_{nz}(\mathbf{r}) & E_{nx}(\mathbf{r}) \end{vmatrix} - \omega_n^2 \mu_{nxx}^{(1)} \varepsilon_{nxx}^{(1)} E_{nx}(\mathbf{r}) = \omega_n^2 \mu_{nxx}^{(1)} P_{nx}^{NL}(\mathbf{r}) \\ & \frac{\partial^2}{\partial x \partial y} E_{ny}(\mathbf{r}) + \frac{\partial^2}{\partial x \partial z} E_{nz}(\mathbf{r}) - \left[\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + \omega_n^2 \mu_{nxx}^{(1)} \varepsilon_{nxx}^{(1)} \right] E_{nx}(\mathbf{r}) = \omega_n^2 \mu_{nxx}^{(1)} P_{nx}^{NL}(\mathbf{r}) \quad (\text{C56}) \end{aligned}$$

在各向同性的电介质中，无论坐标系如何选取，三维二阶张量 $\overrightarrow{\mathbf{I}} \cdot \overrightarrow{\chi}_e^{(1)}(\omega_n)$ 的非主对角元素均全为零，且三个主对角元的值相等，则式(C52)中的 $\overrightarrow{\mathbf{I}} \cdot \overrightarrow{\chi}_e^{(1)}(\omega_n)$ 、 $\overrightarrow{\boldsymbol{\varepsilon}}^{(1)}(\omega_n)$ 全退化为标量 1 、 $\chi_e^{(1)}(\omega_n)$ 、 $\varepsilon_r^{(1)}(\omega_n)$ 、 $\varepsilon^{(1)}(\omega_n)$ ；若同时介质也是磁各向同性的，则此时波动方程(C52)变为：

$$\nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})] - \omega_n^2 \mu^{(1)}(\omega_n) \varepsilon^{(1)}(\omega_n) \mathbf{E}_n(\mathbf{r}) = \omega_n^2 \mu^{(1)}(\omega_n) \mathbf{P}_n^{NL}(\mathbf{r}) \quad (\text{C57})$$

附C1.2.5.2 电矢量是否无源基本只与一阶非线性系数有关

然而即使在电磁各向同性的介质中，若介质同时也是电非线性的 ($\mathbf{P}_n^{NL}(\mathbf{r}) \neq \mathbf{0}$)，则一般也不会有 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) = 0$ ，而只有 $\nabla \cdot \mathbf{D}_n(\mathbf{r}) = 0$ ；即即使式(C52)退化到式(C57)，也无法将其 $\nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})]$ 部分简化；但由于高一阶的非线性系数比相邻低一阶的非线性系数的值，小 12 个数量级^①，则有理由相信式(C56)、(C57)中的非线性波源项

$\mathbf{P}_n^{NL}(\mathbf{r}) \approx \mathbf{0}$ ，因此其解出来的特解 $\mathbf{E}_n(\mathbf{r})$ ，理应类似线性介质中的解(的复振幅)，因此式(C56)、(C57)中的 $\nabla \cdot \mathbf{E}_n(\mathbf{r})$ 是否约等于零，可直接考察无非线性波源项(各分量均为零)的线性介质中的相应波动方程，即电矢量无源与否只需考虑到晶体的一阶非线性系数。所以接下来只需考虑各向同性介质和各向异性介质的电矢量的散度。

下面以电磁各向异性和电磁各向同性的介质为例，验证在 $\mathbf{P}_n^{NL}(\mathbf{r}) \rightarrow \mathbf{0}$ 的极限条件下，对应的解的极限形式是否是无限大平面电磁波的复振幅，且此时解的散度是否为零；其过程只能是自洽式的证明，需先假设大条件 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) = 0$ ，后期望解出来的解又符合或反过来印证此大条件：

以 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) = 0$ 为大前提，代入矢量分析式 $\nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})] = \nabla [\nabla \cdot \mathbf{E}_n(\mathbf{r})] - \nabla^2 \mathbf{E}_n(\mathbf{r}) = -\nabla^2 \mathbf{E}_n(\mathbf{r})$ ，式(C56)、(C57)分别变为：

^① 可将式 1-7 进一步退化为一维标量方程 $\tilde{P}(t) = \varepsilon_0 \sum_{l=1}^{\infty} \chi^{(l)} \tilde{E}^l(t)$ ，并通过量子力学手段，估算出对于凝聚态物质， $\chi^{(1)} \sim 1$ 、 $\chi^{(2)} \approx \chi^{(1)} / E_{at} = 1.94 \times 10^{-12} \text{ m/V}$ ，其中 $E_{at} = e / 4\pi\varepsilon_0 a_0^2 = 5.14 \times 10^{11} \text{ V/m}$ 为特征原子电场强度大小， $a_0 = 4\pi\varepsilon_0 \hbar^2 / m_e e^2$ 为氢原子的玻尔半径；类似地，算得 $\chi^{(3)} \approx 3.78 \times 10^{-24} \text{ m}^2/\text{V}^2$ 。

$$\left[\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial x^2} + \omega_n^2 \mu_{nxx}^{(1)} \varepsilon_{nxx}^{(1)} \right] E_{nx}(\mathbf{r}) = -\omega_n^2 \mu_{nxx}^{(1)} P_{nx}^{NL}(\mathbf{r})$$

$$\nabla^2 \mathbf{E}_n(\mathbf{r}) + \omega_n^2 \mu^{(1)}(\omega_n) \varepsilon^{(1)}(\omega_n) \mathbf{E}_n(\mathbf{r}) = -\omega_n^2 \mu^{(1)}(\omega_n) \mathbf{P}_n^{NL}(\mathbf{r}) \quad (C58)$$

设小前提 $\mathbf{P}_n^{NL}(\mathbf{r}) = \mathbf{0}$ ，则各向异性、同性介质中的定态波动方程变为齐次的——亥姆霍兹方程：

$$\left[\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial x^2} + \omega_n^2 \mu_{nxx}^{(1)} \varepsilon_{nxx}^{(1)} \right] E_{nx}(\mathbf{r}) = 0$$

$$\nabla^2 \mathbf{E}_n(\mathbf{r}) + \omega_n^2 \mu^{(1)}(\omega_n) \varepsilon^{(1)}(\omega_n) \mathbf{E}_n(\mathbf{r}) = \mathbf{0} \quad (C59)$$

附C1.2.5.3 电矢量无源条件下，各向同性介质中波动方程存在有意义的平面波解

可见在无源条件下，电磁各向异性介质的分量方程，已与电磁各向同性介质的分量方程形式一致，因此二者的解的形式相同：在一维/三维直角坐标系下^①，两个方程的特解（不含常数）的形式均要么是三角函数形式，要么是指数为纯虚数、底为自然底数的指数函数形式。现以后一个代表更高维度的矢量式为例，特解为单色平面电磁波的复振幅 $\mathbf{E}_n(\mathbf{r}) = A_n e^{ik_n \cdot r}$ ；乘上分离变量后的时间变量所主导的另一个偏微分方程的特解，即时间项 $e^{-i\omega_n t}$ 得 $\tilde{\mathbf{E}}_n(\mathbf{r}, t) = A_n e^{i(k_n \cdot r - \omega_n t)}$ 。将 $\mathbf{E}_n(\mathbf{r}) = A_n e^{i(k_{nx}x + k_{ny}y + k_{nz}z)}$ 代入方程(C59)有

$$\nabla^2 \mathbf{E}_n(\mathbf{r}) = (-k_{nx}^2 - k_{ny}^2 - k_{nz}^2) \mathbf{E}_n(\mathbf{r}) \text{，而左侧公因式 } \mathbf{E}_n(\mathbf{r}) \neq \mathbf{0} \text{，必有系数部分等于零即}$$

$$\mu^{(1)}(\omega_n) \varepsilon^{(1)}(\omega_n) = 1 / \frac{\omega_n^2}{k_n^2} = 1 / \left(\frac{2\pi/T_n}{2\pi/\lambda_n} \right)^2 = 1 / \left(\frac{\lambda_n}{T_n} \right)^2 = \frac{1}{v_n^2}$$

$$\mu_r^{(1)}(\omega_n) \varepsilon_r^{(1)}(\omega_n) = \frac{c^2}{v_n^2} := n_n^2 \geq 0 \quad (C60)$$

从中可见 $\mu_r^{(1)}(\omega_n) \varepsilon_r^{(1)}(\omega_n)$ 同号，但并非都全为正，可全为负；且折射率 $n_n = \pm \sqrt{\mu_r^{(1)}(\omega_n) \varepsilon_r^{(1)}(\omega_n)}$ 也不一定为正。苏联物理学家 Victor Veselago 在 1968 年指出，

^① 球坐标系和柱坐标系下的解又不同。需根据实际问题的边界形状选择坐标系，而不同坐标系意味着不同边界形状，属于不同的边界条件；而不同的足够的（一般需要两个齐次的）边界条件将约束分离变数后的空间偏微分方程以构成不同的本征值问题（或足够的初始条件与分离变数后的时间偏微分方程构成本征值问题）；相应地可以解出广播天线发出的球面波、波导内定向传播的波、激光器激发的带有束腰的狭窄的光波模式等。这里是空间变量对应的亥姆霍兹方程，但没有边界条件，也就不构成本征值问题，导致波矢不会离散取值，而是连续取值；解为单色平面电磁波。

相对介电常数、相对磁导率、折射率，三者同号。规定其中：

$$|\mathbf{k}_n| = k_n = \frac{\omega_n}{v_n} = \frac{|\omega_n n_n|}{c} = \frac{\omega_n |n_n|}{c}, \text{ 且 } k_n, \omega_n, v_n = \frac{1}{\sqrt{\mu^{(1)}(\omega_n) \epsilon^{(1)}(\omega_n)}} = \frac{c}{|n_n|} \text{ 恒正}^{\circledcirc} \quad (\text{C61})$$

附C1.2.5.4 电矢量无源条件下，电磁均各向异性的介质中波动方程无有意义的解，说明各向异性电磁介质中无法无源近似

而对于电磁各向异性的介质，其 x 分量的方程的解不是 $\mathbf{E}_{nx}(\mathbf{r}) = \mathbf{E}_{nx}(x) = A_{nx} e^{i(k_n x)}$ $= A_{nx} e^{i[(k_n \cdot \mathbf{i}) \cdot \mathbf{r} \cdot \mathbf{i}]}$ ^②，因为物理情景上 $\mathbf{E}_{nx}(\mathbf{r})$ 必须与 $\mathbf{E}_n(\mathbf{r})$ 同相位，因此数学上的解只能为 $\mathbf{E}_{nx}(\mathbf{r}) = A_{nx} e^{ik_n \cdot \mathbf{r}}$ ，但若将此解代入方程(C59)会得 $\mu_{nxx}^{(1)} \epsilon_{nxx}^{(1)} = 1 / (\omega_n^2 / k_n^2) = 1 / v_n^2$ ，与其余两个分量方程所得的 v_n^2 矛盾。之后会知道，这里的电磁均各向异性的波动方程尚不清楚是否有平面电磁波解，所以这里推出矛盾是正常的；只有磁各向同性、电各向异性的介质的波动方程才会有严格的平面电磁波解，并且不需要也不能像这样做电矢量无源近似。因此电磁各向异性的介质内的波动方程只能写作下述电矢量有源形式（毕竟即使有 $\tilde{\rho}_f = 0$ 、 $\nabla \cdot \tilde{\mathbf{D}} = \tilde{\rho}_f$ 、 $\tilde{\mathbf{D}} = \epsilon_0 \tilde{\mathbf{E}} + \tilde{\mathbf{P}}$ ， $\nabla \cdot \tilde{\mathbf{E}} = -\nabla \cdot \tilde{\mathbf{P}} / \epsilon_0 = \tilde{\rho}_p / \epsilon_0$ ，其面内极化/束缚电荷体密度 $\tilde{\rho}_p = -\nabla \cdot \tilde{\mathbf{P}}$ 也不一定为零，因此电矢量不会无源，最多近似无源），且求解时需令其行列式为零：

$$\begin{bmatrix} -\left(\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + \omega_n^2 \mu_{nxx}^{(1)} \epsilon_{nxx}^{(1)}\right) & \frac{\partial^2}{\partial x \partial y} & \frac{\partial^2}{\partial x \partial z} \\ \frac{\partial^2}{\partial y \partial x} & -\left(\frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial x^2} + \omega_n^2 \mu_{nyy}^{(1)} \epsilon_{nyy}^{(1)}\right) & \frac{\partial^2}{\partial y \partial z} \\ \frac{\partial^2}{\partial z \partial x} & \frac{\partial^2}{\partial z \partial y} & -\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \omega_n^2 \mu_{nzz}^{(1)} \epsilon_{nzz}^{(1)}\right) \end{bmatrix} \begin{bmatrix} E_{nx} \\ E_{ny} \\ E_{nz} \end{bmatrix}$$

^①之所以定义 $k_n > 0$ ，是方便配合 $\hat{\mathbf{k}}_n$ 描述 \mathbf{k}_n ，使得 $\mathbf{k}_n = k_n \hat{\mathbf{k}}_n$ 与 $\hat{\mathbf{k}}_n$ 恒同向；同理这也是定义 $v_n > 0$ 的原因。之前确实定义过负频率。但负频率同时也对应着负波矢，再取复共轭，则又回到 ω_n, \mathbf{k}_n 双正了， $\therefore \omega_n > 0$ 。

^②这意味着就不会有 $v_n = \sqrt{v_{nx}^2 + v_{ny}^2 + v_{nz}^2} = \sqrt{\frac{1}{\mu_{nxx}^{(1)} \epsilon_{nxx}^{(1)}} + \frac{1}{\mu_{nyy}^{(1)} \epsilon_{nyy}^{(1)}} + \frac{1}{\mu_{nzz}^{(1)} \epsilon_{nzz}^{(1)}}} = \frac{c}{|n_n|}$ ，以及

$\frac{1}{n_n^2} = \frac{1}{\mu_{rn}^{(1)} \epsilon_{rn}^{(1)}} = \frac{1}{\mu_{rnxx}^{(1)} \epsilon_{rnxx}^{(1)}} + \frac{1}{\mu_{rnyy}^{(1)} \epsilon_{rnyy}^{(1)}} + \frac{1}{\mu_{rnzz}^{(1)} \epsilon_{rnzz}^{(1)}} = \frac{1}{n_{nx}^2} + \frac{1}{n_{ny}^2} + \frac{1}{n_{nz}^2}$ ；并说明 $v_{nx}^2 = \frac{1}{\mu_{nxx}^{(1)} \epsilon_{nxx}^{(1)}}$ 等

三者均不成立。

$$= \omega_n^2 \begin{bmatrix} \mu_{nxx}^{(1)} P_{nx}^{NL} \\ \mu_{nyy}^{(1)} P_{ny}^{NL} \\ \mu_{nzz}^{(1)} P_{nz}^{NL} \end{bmatrix} \quad (C62)$$

附C1.2.5.5 定义各向异性介质中的折射率椭球，为下文做铺垫

为与之后的推导相适应，现定义对于电磁各向异性介质的折射率椭球为：

$$\mathbf{n}_n^T \cdot \overrightarrow{\boldsymbol{\mu}}_{rn}^{(1)} \cdot \overrightarrow{\boldsymbol{\epsilon}}_{rn}^{(1)} \cdot \mathbf{n}_n = (\mathbf{n}_n \mathbf{I}) \cdot \overrightarrow{\boldsymbol{\mu}}_{rn}^{(1)} \cdot \overrightarrow{\boldsymbol{\epsilon}}_{rn}^{(1)} \cdot (\mathbf{n}_n \mathbf{I}) = \frac{(n_n x_0)^2}{n_{nx}^2} + \frac{(n_n y_0)^2}{n_{ny}^2} + \frac{(n_n z_0)^2}{n_{nz}^2} = 1$$

$$\frac{x_0^2}{n_{nx}^2} + \frac{y_0^2}{n_{ny}^2} + \frac{z_0^2}{n_{nz}^2} = \frac{1}{n_n^2} \quad (C63)$$

其中 $\mathbf{n}_n = n_n \mathbf{I}$ 为在主轴化后的坐标系下，所考虑的某方向的折射率，

$\mathbf{I} = x_0 \mathbf{i} + y_0 \mathbf{j} + z_0 \mathbf{k}$ 为该方向的单位矢量；其中也有类似的 $n_{nx} = \pm \sqrt{\mu_{rnx}^{(1)} \epsilon_{rnx}^{(1)}}$ ，且 $n_{nx}, \mu_{rnx}^{(1)}, \epsilon_{rnx}^{(1)}$ 三者同号；以及可定义

$$\frac{1}{n_n} := \pm \sqrt{\frac{x_0^2}{n_{nx}^2} + \frac{y_0^2}{n_{ny}^2} + \frac{z_0^2}{n_{nz}^2}} \quad (C64)$$

其中 $n_{nx}, n_{ny}, n_{nz}, n_n, \mu_{rn}^{(1)}, \epsilon_{rn}^{(1)}$ 六者同号；由此可见并没有

$$\frac{1}{n_n^2} = \frac{1}{n_{nx}^2} + \frac{1}{n_{ny}^2} + \frac{1}{n_{nz}^2} \quad (C65)$$

折射率椭球所描绘的 n_n 实际上是一种“等效于折射率张量的，折射率标量”，其牺牲了阶数维数而换取了成为方向的标量函数，可用于考察 $\mathbf{P}, \mathbf{D}, \mathbf{M}, \mathbf{B}$ 的偏振方向上的等效折射率 n_n 的大小，进而在之后可用于确定诸如磁各向同性的线性介质中的 $\mathbf{D}_n(\mathbf{r}) = \epsilon_0 n_{ne} |E_{n\perp}(\mathbf{r})|$ 的大小，而不需要采用张量的形式 $\tilde{\mathbf{D}}_n = \overrightarrow{\boldsymbol{\epsilon}}_n^{(1)} \cdot \tilde{\mathbf{E}}_n$ ；折射率椭球按理说也可用于确定波矢方向上的等效折射率 n_n 的大小，进而考察相速度、波矢的大小，但之后会发现这由折射率曲面负责，比如折射率椭球可描绘 e 光但不可描绘 o 光的相速度曲面和波矢曲面，但本质上折射率椭球可导出折射率曲面，且 o 光的 $\mathbf{P}, \mathbf{D}, \mathbf{M}, \mathbf{B}$ 也将与 e 光的极化程度一起由折射率椭球决定。

附C1.2.6 无源近似下单色平面波解的物理意义、复色场群速度与相速度的关系

附C1.2.6.1 电矢量无源近似下，各向同性介质中平面波的物理意义

考察解 $\tilde{\mathbf{E}}_n$ 的相位部分 $\mathbf{k}_n \cdot \mathbf{r} - \omega_n t = C$ 为一常数并对两边微分得：

$$\nu_n = \frac{\omega_n}{k_n} = \frac{dr_{\hat{k}_n}}{dt} = \frac{d(r \cdot \cos\langle \mathbf{k}_n, \mathbf{r} \rangle)}{dt} \quad (C66)$$

其中规定 $\hat{\nu}_n = \hat{\mathbf{k}}_n$ ，则因 ν_n 、 k_n 均为正，有相速度 ν_n 与波矢 \mathbf{k}_n 同向，则 $\hat{\nu}_n, \hat{\mathbf{k}}_n$ 均可作为波面的外法向单位矢量、均可代表波的传播方向，称作波法线。而相速度

$$\mathbf{v}_n = \nu_n \hat{\mathbf{v}}_n = \frac{d(r \cdot \cos\langle \mathbf{k}_n, \mathbf{r} \rangle)}{dt} \hat{\mathbf{k}}_n = \frac{d(r \cdot \cos\langle \mathbf{v}_n, \mathbf{r} \rangle)}{dt} \hat{\mathbf{v}}_n \quad (C67)$$

物理意义为等相位面的传播速度。由于任一时间断面下，解 $\tilde{\mathbf{E}}_n(\mathbf{r}, t) = A_n e^{i(k_n \cdot \mathbf{r} - \omega_n t)}$ 在三维空间中的等相位面均为一平面，形状不随时间改变，所以称其为单色平面电磁波，其等相位面即波(阵)面上各子球面波的传播速度大小，和包络面构成的各面元外法向矢量的方向，均不随空间和时间变化，这由相位部分的空间(含 \mathbf{r})部分只有 $\mathbf{k}_n \cdot \mathbf{r}$ ，且其中的 \mathbf{k}_n 是一不含时也不含空的常矢量决定。波阵面各处均以等大同向的不变速度矢量 \mathbf{v}_n 向前传播，这就抽象出来了相速度，可见相速度还是比较“宏观的”，它统治了一个点波源群；而从平面波的波阵面也是无数子球面波的包络面的角度，以及平面波的等相位面上各点的电矢量瞬时值 $\tilde{\mathbf{E}}_n$ 相等的角度，相速度也是特殊的群速度。

附C1.2.6.2 两单色场叠加而成的复色场的振幅波与群速度、相位波与相速度

群速度相比于相速度，相通之处是群速度恒考察包络面的传播速度，相速度只在波为单色平面波的时候才考察包络面的传播速度；不同之处是 ν_g 考察的是固定振幅(最大瞬时值) $|\tilde{\mathbf{E}}| = A$ 的点的传播速度，而不考察固定相位 $\tilde{\mathbf{E}}/|\tilde{\mathbf{E}}| = \tilde{\mathbf{E}}/A$ 的点的传播速度(相速度)，也不考虑固定瞬时值 $\tilde{\mathbf{E}}$ 的点的传播速度(因为对于一般的非行波的复合场，没法考虑这一点)。从两束偏振方向和振幅均相同、频率波矢均不相同但相差甚小(甚大也可以，甚至波矢方向不同也可)的单色波的叠加场 $\tilde{\mathbf{E}}_1 + \tilde{\mathbf{E}}_2 = Ae^{i(k_1 \cdot \mathbf{r} - \omega_1 t)} + Ae^{i(k_2 \cdot \mathbf{r} - \omega_2 t)} = Ae^{i(\frac{k_1 - k_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2}t)} e^{i(\frac{k_1 + k_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2}t)} + Ae^{-i(\frac{k_1 - k_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2}t)} e^{i(\frac{k_1 + k_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2}t)} = 2A \cos(\frac{k_1 - k_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2}t) \cdot e^{i(\frac{k_1 + k_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2}t)}$ 可看出，振幅部分是个“包裹着”相位部分单色波的包络波：

当两个子波满足 $\mathbf{k}_1 \cdot \mathbf{k}_2 > 0$ 时，合成波的包络波的波矢 $|(\mathbf{k}_1 - \mathbf{k}_2)/2| < |(\mathbf{k}_1 + \mathbf{k}_2)/2|$ ；若进一步在 $\mathbf{k}_1, \mathbf{k}_2$ 同向即 $\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2 = 1$ 时， $(\mathbf{k}_1 - \mathbf{k}_2)/2$ 甚至可以为负，相应的 $(\mathbf{k}_1 - \mathbf{k}_2)/2$ 与被包络的单色波的波矢 $(\mathbf{k}_1 + \mathbf{k}_2)/2$ 反向，但此时并不意味着包络波的传播方向与被包络的相位波相反，只有当 $(\omega_1 - \omega_2)/2 > 0$ 同时也被满足时，才有此结论，因为此时包络波的波矢并不代表其传播方向(因为包络波的频率与波矢可以异号；相位波就不存在此问题)，要波速才代表；由下一段知，普遍结论是 $(\omega_1 - \omega_2)(k_1 - k_2) < 0$ 为群速度与相速度

相反的充要条件。

附C1.2.6.3 振幅/包络波的相速度（即复色场的群速度）与其波矢、与相位波/复色场的相速度均不一定同向

设 $\Delta\omega, \Delta\mathbf{k} = 2\mathbf{k}_g, \Delta k$ ^① 分别代表 $\omega_1 - \omega_2, \mathbf{k}_1 - \mathbf{k}_2, k_1 - k_2$ 或 $\omega_2 - \omega_1, \mathbf{k}_2 - \mathbf{k}_1, k_2 - k_1$ ，易得包络波的其余有关各量：波长 $\lambda_g = 2\pi/|(\mathbf{k}_1 - \mathbf{k}_2)/2| = 4\pi/|\mathbf{k}_1 - \mathbf{k}_2| = 4\pi/|\Delta\mathbf{k}|$ ，在 $\langle \hat{\mathbf{k}}_1, \hat{\mathbf{k}}_2 \rangle < \pi/2$ 的条件下， $\lambda_g > \lambda_v$ ，意味着包络波的空间周期一般长于相位波，印证了振幅波“包络、调制”相位波的结论；令振幅波相位部分为常数并微分，可赋予群速度物理意义。现定义群速度

$$\begin{aligned} \mathbf{v}_g &:= \frac{d(r \cdot \cos\langle \mathbf{k}_g, \mathbf{r} \rangle)}{dt} \hat{\mathbf{k}}_g = \frac{dr_{\hat{\mathbf{k}}_g}}{dt} \hat{\mathbf{k}}_g = \frac{\frac{\omega_1 - \omega_2}{2}}{\left|\frac{\mathbf{k}_1 - \mathbf{k}_2}{2}\right|} \left(\frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \right) = \frac{\omega_1 - \omega_2}{|\mathbf{k}_1 - \mathbf{k}_2|} \hat{\mathbf{k}}_1 - \hat{\mathbf{k}}_2 \\ &= \frac{\Delta\omega}{|\Delta\mathbf{k}|} \hat{\Delta\mathbf{k}} = \frac{\omega_g}{k_g} \hat{\mathbf{k}}_g := \frac{c}{n_g} \operatorname{sgn}(\hat{\mathbf{v}}_g \cdot \hat{\mathbf{v}}_p) \hat{\mathbf{v}}_g \end{aligned} \quad (\text{C68})$$

其中 $k_g = |\Delta\mathbf{k}/2| > 0$ ，且

$$v_g = \frac{|\omega_g|}{k_g} = \frac{dr_{\hat{\mathbf{k}}_g}}{dt} = \frac{d(r \cdot \cos\langle \mathbf{k}_g, \mathbf{r} \rangle)}{dt} = \frac{c}{n_g} \operatorname{sgn}(\hat{\mathbf{v}}_g \cdot \hat{\mathbf{v}}_p) > 0 \quad (\text{C69})$$

其中 $\Delta\omega, \hat{\Delta\mathbf{k}}, \hat{\mathbf{k}}_g, n_g, \hat{\mathbf{v}}_g \cdot \hat{\mathbf{v}}_p$ 均可为负。

群速度表达式两侧点乘 $\hat{\mathbf{k}}_g$ ，可得到之前的结论的拓展版：

$$\hat{\mathbf{v}}_g \cdot \hat{\mathbf{k}}_g = \frac{v_g k_g}{\omega_g} = \frac{|\omega_g|}{\omega_g} = \frac{\Delta\omega}{|\Delta\omega|} \quad (\text{C70})$$

相比 $\hat{\mathbf{k}}_p \cdot \hat{\mathbf{v}}_p = v_p k_p / \omega_p > 0$ ，包络波的波矢与其相速度不一定同向；又因 $\hat{\mathbf{k}}_g \cdot \hat{\mathbf{k}}_p$ 同号于： $\mathbf{k}_g \cdot \mathbf{k}_p = (\mathbf{k}_1 - \mathbf{k}_2)/2 \cdot (\mathbf{k}_1 + \mathbf{k}_2)/2 = (\mathbf{k}_1^2 - \mathbf{k}_2^2)/4$ 同号于 $k_1 - k_2$ ，则 $\hat{\mathbf{v}}_g \cdot \hat{\mathbf{v}}_p$ 同号于 $(\hat{\mathbf{v}}_g \cdot \hat{\mathbf{k}}_g)(\hat{\mathbf{k}}_g \cdot \hat{\mathbf{k}}_p)(\hat{\mathbf{k}}_p \cdot \hat{\mathbf{v}}_p)$ 同号于 $(\omega_1 - \omega_2)(k_1 - k_2)v_p k_p / \omega_p$ 同号于 $\Delta\omega \cdot \Delta k$ ；又 c/n_g 同号于 $\hat{\mathbf{v}}_g \cdot \hat{\mathbf{v}}_p$ ，所以 c/n_g 也同号于 $\Delta\omega \cdot \Delta k$ ，即有 n_g 同号于 $\Delta\omega \cdot \Delta k$ 。对于定振幅的单色平面波而言，其振幅的平方是正比于其平均能流 $\bar{\mathbf{S}}$ 的，而包络波就是叠加场的振幅部分，虽然它不是定振幅且与时空有关，是个行波，但它作为振幅，其平方的平均值即使不与

^① $\Delta k := k_1 - k_2 \neq |\mathbf{k}_1 - \mathbf{k}_2| = |\Delta\mathbf{k}|$

平均能流成正比，至少也应成正相关关系。因此接下来考察一下相位中含有群速度的包络波，其本身的传播方向，即群速度的方向，是否与叠加场的能流方向相同，或者有关。

附C1.2.7 复色场能流密度表明必须考虑负频光和差频光，其与群速度无显著关联

现暂将 $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ 的振幅部分写作 $\mathbf{A} = \mathbf{E}_{01} = \mathbf{E}_{02} = \mathbf{E}_0$ ^①，仿照 $\tilde{\mathbf{E}} = \tilde{\mathbf{E}}_1 + \tilde{\mathbf{E}}_2 = 2\mathbf{E}_0 \cos(\frac{k_1-k_2}{2} \cdot \mathbf{r} - \frac{\omega_1-\omega_2}{2}t) \cdot e^{i(\frac{k_1+k_2}{2} \cdot \mathbf{r} - \frac{\omega_1+\omega_2}{2}t)}$ 的得来过程，只有当 $\tilde{\mathbf{H}}_1, \tilde{\mathbf{H}}_2$ 也同向等大时，才有 $\tilde{\mathbf{H}} = \tilde{\mathbf{H}}_1 + \tilde{\mathbf{H}}_2 = 2\mathbf{H}_0 \cos(\frac{k_1-k_2}{2} \cdot \mathbf{r} - \frac{\omega_1-\omega_2}{2}t) \cdot e^{i(\frac{k_1+k_2}{2} \cdot \mathbf{r} - \frac{\omega_1+\omega_2}{2}t)}$ ，此时叠加磁场的包络波的相位部分和相位波才分别与电场的相同。考虑 $\langle \mathbf{S}_1(\mathbf{r}), \hat{\mathbf{k}}_1 \rangle = \langle \mathbf{S}_2(\mathbf{r}), \hat{\mathbf{k}}_2 \rangle = 0$ ，即两束光都均匀极化（如都处在各向同性介质中）、二者的离散角 $\langle \mathbf{E}_i(\mathbf{r}), \mathbf{D}_i(\mathbf{r}) \rangle = \alpha_i = 0$ 相同，那么此时由于两束偏振光的 $\mathbf{E}_1, \mathbf{E}_2, \mathbf{H}_1, \mathbf{H}_2$ 均分别等大同向，则两束光的 $\mathbf{S}_1(\mathbf{r}), \mathbf{S}_2(\mathbf{r})$ 也分别等大同向，又因二者的离散角均为 0，则两束光的 $\hat{\mathbf{k}}_1, \hat{\mathbf{k}}_2$ 也分别同向。再根据 $\mathbf{S}_1(\mathbf{r}), \mathbf{S}_2(\mathbf{r})$ 等大，可得 $\mathbf{S}_i(\mathbf{r}) = w_i \mathbf{v}_{ir} = w_i \mathbf{v}_i = (\mathbf{E}_i \cdot \mathbf{D}_i) \mathbf{v}_i = \varepsilon_i^{(1)} E_i^2 \mathbf{v}_i$ 等大，而又因 $\mathbf{E}_1, \mathbf{E}_2$ 等大，有 $\varepsilon_i^{(1)} v_i$ 相等，同理 $\mu_i^{(1)} v_i$ 相等；而如果同一晶体中两束光的光路上的 $\varepsilon_i^{(1)}$ 相同，则两束光的 $v_i = \omega_i/k_i$ 相同，且两束光光路上的 $\mu_i^{(1)}$ 相同；但此时还剩下一个自由度，即两束光虽有相速度相同，频率和波矢却可以互不相同，不过需要成比例，因此此时两束光还不一定是完全一样的两束光，所得的

$$\tilde{\mathbf{S}} = \tilde{\mathbf{E}} \times \tilde{\mathbf{H}} = 4\mathbf{S}_0 \cos^2(\frac{k_1-k_2}{2} \cdot \mathbf{r} - \frac{\omega_1-\omega_2}{2}t) \cdot e^{i[(k_1+k_2) \cdot \mathbf{r} - (\omega_1+\omega_2)t]} \quad (C71)$$

还是有意义的，此时 $\hat{\mathbf{k}}_1, \hat{\mathbf{k}}_2, \hat{\mathbf{k}}_p$ 共线同向， $\hat{\mathbf{v}}_g \cdot \hat{\mathbf{v}}_p$ 与 $\Delta\omega \cdot \Delta k$ 均为正，且 $\hat{\mathbf{v}}_g, \hat{\mathbf{v}}_p$ 同向，即 $\hat{\mathbf{v}}_g \cdot \hat{\mathbf{v}}_p = 1$ ；且 $\hat{\mathbf{S}} \cdot \hat{\mathbf{v}}_p = \hat{\mathbf{S}} \cdot \hat{\mathbf{k}}_p = 1$ 。因此此时 $\hat{\mathbf{S}}$ 与 $\hat{\mathbf{v}}_g$ 同向，即 $\hat{\mathbf{S}} \cdot \hat{\mathbf{v}}_g = 1$ 。

附C1.2.7.1 能流是可观测量，必须是实的，由此导出单色能流及其时空平均值表达式

但上述过程忽略了一个事实：由于能流是可测量的物理量，因此它必须是实数，一种办法是对其取模，但这就失去了其相位部分和其中与时空相关的特性，只剩下振

^① 下标含 0 且 0 在 n 或 i 前的标量/矢量，在本文中表示定振幅的标量/矢量；下标含 0 且 0 在 n 或 i 后的标量/矢量，表示定振幅的单位标量/矢量。如

$$\hat{\tilde{\mathbf{E}}}_n = \hat{\mathbf{E}}_n = \mathbf{E}_{n0} = \frac{\mathbf{E}_{0n}}{|\mathbf{E}_{0n}|} = \frac{\mathbf{E}_{0n}}{0.5 \cdot \mathcal{E}_n} = \frac{0.5 \cdot \mathcal{E}_n}{\mathcal{E}_n} = \frac{\mathcal{E}_n}{\mathcal{E}_n}；一般不能含 0 的同时而不含 n，否则表示展$$

开式中的直流分量成分，即认为 n=0，但这里是例外。

幅部分即包络波。因此正确的做法和更常见的是，对于 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_i (i \geq 0)$ 频率、波矢均等的单色波，有

$$\begin{aligned}\tilde{\mathbf{S}}_i &= \tilde{\mathcal{E}}_i \times \tilde{\mathcal{H}}_i = (\tilde{\mathbf{E}}_i + \tilde{\mathbf{E}}_i^*) \times (\tilde{\mathbf{H}}_i + \tilde{\mathbf{H}}_i^*) = \mathcal{E}_i \cos(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t) \times \mathcal{H}_i \cos(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t) \\ &= \mathcal{S}_i \cos^2(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t) = \mathcal{S}_i \frac{1 + \cos(2\mathbf{k}_i \cdot \mathbf{r} - 2\omega_i t)}{2}\end{aligned}\quad (C72)$$

当 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_i$ 频率不同时，不论波矢相同与否， $\tilde{\mathbf{S}}_i$ 的时间平均值在空间各处均为零；当 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_i$ 波矢不同、频率相同时， $\tilde{\mathbf{S}}_i$ 的时间平均值虽不为零，但在固定空间点上可能为某一正的或负的常值，且该常值虽与时间无关，但仍随空间起伏，且有正负；只有当 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_i$ 的波矢、频率均相同时， $\tilde{\mathbf{S}}_i$ 的时间平均值在空间各处的常值才恒正，且与空间无关。——但即使 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_i (i \neq 0)$ 的频率、波矢均不同，其中

$$\begin{aligned}\tilde{\mathbf{E}}_i \times \tilde{\mathbf{H}}_i + \tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i^* &= \mathcal{S}_{0i} e^{i(2\mathbf{k}_i \cdot \mathbf{r} - 2\omega_i t)} + \mathcal{S}_{0i} e^{-i(2\mathbf{k}_i \cdot \mathbf{r} - 2\omega_i t)} = \tilde{\mathbf{E}}_i \times \tilde{\mathbf{H}}_i + (\tilde{\mathbf{E}}_i \times \tilde{\mathbf{H}}_i)^* \\ &= 2\mathcal{S}_{0i} \cos(2\mathbf{k}_i \cdot \mathbf{r} - 2\omega_i t) = \frac{\mathcal{S}_i}{2} \cos(2\mathbf{k}_i \cdot \mathbf{r} - 2\omega_i t) (i \neq 0)\end{aligned}\quad (C73)$$

部分在时间或空间上的平均值无论如何也都为零^②。因此对 $\tilde{\mathbf{S}}_i$ 取时间平均值即有

$$\mathbf{I}_i = \overline{\tilde{\mathbf{S}}_i} = \overline{\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i + \tilde{\mathbf{E}}_i \times \tilde{\mathbf{H}}_i^*} = \overline{\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i + (\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i)^*} (i \neq 0) \quad (C74)$$

同样的原因，当 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_i$ 频率相同时，即使波矢不同， $\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i + (\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i)^*$ 中虽有 cos 因子但 cos 中含空不含时，所以它不含时，此时其时间平均值可去掉时间平均符号以得到：

$$\mathbf{I}_i = \tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i + (\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i)^* = \mathbf{E}_i^* \times \mathbf{H}_i + (\mathbf{E}_i^* \times \mathbf{H}_i)^* (i \neq 0) \quad (C75)$$

但它可能含空。只有当 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_i$ 频率和波矢都相同时，才有 \mathbf{I}_i 为一不含时也不含空

^① $\tilde{\mathbf{S}}_n = \tilde{\mathbf{E}}_n \times \tilde{\mathbf{H}}_n = \mathcal{S}_n(\mathbf{r}) e^{-i2\omega_n t}$ 一般不成立，而没有所谓的 $\mathcal{S}_n(\mathbf{r})$ ；所以对于既含时又含空的 $\tilde{\mathbf{S}}_n$ ，其时空变量在 cos 的相位中无法分离；除非 $\tilde{\mathbf{S}}_n$ 只含时或只含空，但此时谈分离又没有意义了。

^② 这是因为其 cos 内波矢和频率为两非零正值之和，而都无法为零，所以不仅恒有 cos 因子，而且 cos 内恒含时且含空；注意这里的讨论默认了 $i \geq 1$ ，当 $i = 0$ 时 $\tilde{\mathbf{S}}_i$ 与 \mathbf{I}_i 中都应考虑并加上此时的 $\tilde{\mathbf{E}}_i \times \tilde{\mathbf{H}}_i + \tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i^* = \frac{\mathcal{S}_i}{2} \cos(2\mathbf{k}_i \cdot \mathbf{r} - 2\omega_i t) = \frac{\mathcal{S}_i}{2}$ 。

的常值，此时：

$$\mathbf{I}_i = 2\mathbf{S}_{0i} = \frac{\mathcal{S}_i}{2} (i \neq 0), \quad \mathbf{I}_0 = 4\mathbf{S}_{00} = \mathcal{S}_0 \quad (\text{C76})$$

附C1.2.7.2 复色场的能流及其时间平均值

(1) 由复单色子波构成的复复色场的能流

若叠加场 $\tilde{\mathbf{E}} = \sum_{i=0}^{\infty} \tilde{\mathbf{E}}_i$ 与 $\tilde{\mathbf{H}} = \sum_{i=0}^{\infty} \tilde{\mathbf{H}}_i$ 的相同角标的子波均来自同一条电磁波，则 $\tilde{\mathbf{S}} = \tilde{\mathbf{E}} \times \tilde{\mathbf{H}} = (\tilde{\mathbf{E}} + \tilde{\mathbf{E}}^*) \times (\tilde{\mathbf{H}} + \tilde{\mathbf{H}}^*)$ ，同样无论任意一对 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_j (i, j \neq 0)$ 的波矢、频率是否相同，均有：

$$\begin{aligned} \tilde{\mathbf{E}} \times \tilde{\mathbf{H}} + (\tilde{\mathbf{E}} \times \tilde{\mathbf{H}})^* &= \sum_{i=0}^{\infty} \tilde{\mathbf{E}}_i \times \sum_{i=0}^{\infty} \tilde{\mathbf{H}}_i + \left(\sum_{i=0}^{\infty} \tilde{\mathbf{E}}_i \times \sum_{i=0}^{\infty} \tilde{\mathbf{H}}_i \right)^* \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \left[\tilde{\mathbf{E}}_i \times \tilde{\mathbf{H}}_j + (\tilde{\mathbf{E}}_i \times \tilde{\mathbf{H}}_j)^* \right] (|i| + |j| \neq 0) \end{aligned} \quad (\text{C77})$$

的时间、空间平均值全等于零。那么仍有：

$$\overline{\mathbf{I}} = \overline{\tilde{\mathbf{E}}^* \times \tilde{\mathbf{H}} + (\tilde{\mathbf{E}}^* \times \tilde{\mathbf{H}})^*} = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \overline{\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_j + (\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_j)^*} (|i| + |j| \neq 0) \quad (\text{C78})$$

其中若角标取值不同的 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_j (i \neq j)$ 的频率不同，则相应 $\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_j + (\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_j)^*$ 的时间平均值等于零，此时(时间)平均能流密度化为：

$$\overline{\mathbf{I}} = \sum_{i=0}^{\infty} \overline{\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i + (\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i)^*} (i \neq 0) = \sum_{i=1}^{\infty} \overline{\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i + (\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i)^*} \quad (\text{C79})$$

而一般 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_i$ 频率相同（否则其 $\mathbf{I}_i = \mathbf{0}$ ，这样的波也就不存在了），因此有包含和不包含背景直流分量 $\mathbf{I}_0 = \overline{\tilde{\mathbf{S}}_0} = \mathcal{S}_0$ 的能流密度时间平均值：

$$\begin{aligned} \overline{\mathbf{I}} &= \sum_{i=1}^{\infty} \left[\overline{\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i + (\tilde{\mathbf{E}}_i^* \times \tilde{\mathbf{H}}_i)^*} \right] = \sum_{i=1}^{\infty} \left[\overline{\mathbf{E}_i^* \times \mathbf{H}_i + (\mathbf{E}_i^* \times \mathbf{H}_i)^*} \right] = \sum_{i=1}^{\infty} \mathbf{I}_i \\ \overline{\mathbf{I}} &= \sum_{i=1}^{\infty} \mathbf{I}_i + \mathbf{I}_0 = \sum_{i=0}^{\infty} \mathbf{I}_i \end{aligned} \quad (\text{C80})$$

一般 $\tilde{\mathbf{E}}_i, \tilde{\mathbf{H}}_i$ 波矢也相同（否则其 \mathbf{I}_i 的 \cos 中含波矢且含空），因此 $\overline{\mathbf{I}}$ 不含时也不含空。但不像 $\overline{\mathbf{I}} = \sum_{i=0}^{\infty} \mathbf{I}_i$ ，含时(瞬时)能流密度本身不遵循 $\tilde{\mathbf{S}} = \sum_{i=0}^{\infty} \tilde{\mathbf{S}}_i$ ，只有其时间平均值满足 $\overline{\tilde{\mathbf{S}}} = \sum_{i=0}^{\infty} \overline{\tilde{\mathbf{S}}_i}$ ，同理其空间平均值也满足类似关系式。

(2) 由实单色子波，或包含负频的成对复单色子波构成的实复色场的能流

若所考虑的叠加场对于所有正频子波已包含其相应负频如 $\tilde{\mathbf{E}} = \sum_{i=-\infty}^{\infty} \tilde{\mathbf{E}}_i = \sum_{i=0}^{\infty} \tilde{\mathbf{E}}_i$ 与 $\tilde{\mathbf{H}} = \sum_{i=-\infty}^{\infty} \tilde{\mathbf{H}}_i = \sum_{i=0}^{\infty} \tilde{\mathbf{H}}_i$ ，则叠加场已经是实的了，其能流密度直接满足 $\tilde{\mathbf{S}} = \tilde{\mathbf{E}} \times \tilde{\mathbf{H}}$ ，

不需要也不能有 $\tilde{\mathbf{S}} = (\tilde{\mathbf{E}} + \tilde{\mathbf{E}}^*) \times (\tilde{\mathbf{H}} + \tilde{\mathbf{H}}^*)$ ，但结论仍可以沿用 $\mathbf{I} = \sum_{i=0}^{\infty} \mathbf{I}_i$ 、 $\bar{\mathbf{S}} = \sum_{i=0}^{\infty} \bar{\mathbf{S}}_i$ ，正如复复色场也可以用 $\tilde{\mathbf{S}} = \tilde{\mathbf{E}} \times \tilde{\mathbf{H}}$ 一样。——从这里对能流密度的分析可看出，在将单色波从实的 \cos 形式写为复数形式是必要的同时，即对叠加场进行复数形式的傅里叶展开以方便时间项分离的同时，负频和负波矢的存在也将与之同等必要，因为需要用之以分离时间项求解后合并时间项，再将复数形式的正负频电磁波解转为实数 \cos 形式，赋予其物理意义，特别是继而赋予类似能流密度等可观测量的物理意义；因此式(C1)与式(C18)在以复数形式进行非线性过程计算的同时，需同时考虑正频与负频的复数形式的非线性过程，且在和频光之外，确实还有差频光，甚至纯和频的负频光的存在。

附C1.2.7.3 复色场能流密度与群速度通常不同向，且值的大小无关联

因此当 $\tilde{\mathbf{H}}_1, \tilde{\mathbf{H}}_2$ 也同向等大时，之前的 $\tilde{\mathbf{S}} = \tilde{\mathbf{E}} \times \tilde{\mathbf{H}}$ 应改为：

$$\begin{aligned}\tilde{\mathbf{S}} &= (\tilde{\mathbf{E}} + \tilde{\mathbf{E}}^*) \times (\tilde{\mathbf{H}} + \tilde{\mathbf{H}}^*) = 4\mathcal{S}_0 \cos^2\left(\frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2}t\right) \cdot \cos^2\left(\frac{\mathbf{k}_1 + \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2}t\right) \\ &= \mathcal{S}_0 [1 + \cos(\Delta\mathbf{k} \cdot \mathbf{r} - \Delta\omega \cdot t)] \cdot \{1 + \cos[(\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r} - (\omega_1 + \omega_2) \cdot t]\}\end{aligned}\quad (C81)$$

其中余弦的二次项由于 $\Delta\omega \neq \omega_1 + \omega_2$ ，对时间取平均值后为零；同理两个余弦的一次项因 $\Delta\omega, \omega_1 + \omega_2 \neq 0$ ，对时间取平均值后也均为零。因此 $\mathbf{I} = \mathcal{S}_0 = 4\mathcal{S}_0$ ，与上上一段的结论 $\mathbf{I} = \sum_{i=1}^2 \mathbf{I}_i = \sum_{i=1}^2 \bar{\mathbf{S}}_i = \sum_{i=1}^2 [\mathbf{E}_i^* \times \mathbf{H}_i + (\mathbf{E}_i^* \times \mathbf{H}_i)^*] = 2\mathcal{S}_0 + 2\mathcal{S}_0 = 4\mathcal{S}_0$ 结果一致；且同样有 $\hat{\mathbf{S}} \cdot \hat{\mathbf{v}}_g = 1$ ，并且由于 $\tilde{\mathbf{S}}$ 与 \mathcal{S}_0 恒同向，因此 $\tilde{\mathbf{S}}$ 也恒与 \mathbf{v}_g 同向，这与 $\tilde{\mathbf{E}} + \tilde{\mathbf{E}}^*$ 和 $\tilde{\mathbf{H}} + \tilde{\mathbf{H}}^*$ 随其相位同时取正负相协调。

对于 $\tilde{\mathbf{H}}_1, \tilde{\mathbf{H}}_2$ 不等大同向的一般情况，

$$\begin{aligned}\tilde{\mathbf{S}} &= (\tilde{\mathbf{E}} + \tilde{\mathbf{E}}^*) \times (\tilde{\mathbf{H}} + \tilde{\mathbf{H}}^*) \\ &= 4\mathbf{E}_0 \cos\left(\frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2}t\right) \cdot \cos\left(\frac{\mathbf{k}_1 + \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2}t\right) \times [(\tilde{\mathbf{H}}_1 + \tilde{\mathbf{H}}_1^*) + (\tilde{\mathbf{H}}_2 + \tilde{\mathbf{H}}_2^*)] \\ &= \frac{4\mathbf{E}_0 \cos\left(\frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2}t\right) \cdot \cos\left(\frac{\mathbf{k}_1 + \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2}t\right)}{2\mathbf{E}_0 \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega_1 t)} (\tilde{\mathbf{E}}_1 + \tilde{\mathbf{E}}_1^*) \times (\tilde{\mathbf{H}}_1 + \tilde{\mathbf{H}}_1^*) \\ &\quad + \frac{4\mathbf{E}_0 \cos\left(\frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2}t\right) \cdot \cos\left(\frac{\mathbf{k}_1 + \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2}t\right)}{2\mathbf{E}_0 \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega_2 t)} (\tilde{\mathbf{E}}_2 + \tilde{\mathbf{E}}_2^*) \times (\tilde{\mathbf{H}}_2 + \tilde{\mathbf{H}}_2^*) \\ &= 2 \left[\frac{\cos\left(\frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2}t\right) \cdot \cos\left(\frac{\mathbf{k}_1 + \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2}t\right)}{\cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega_1 t)} \tilde{\mathbf{S}}_1 + \frac{\cos\left(\frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2}t\right) \cdot \cos\left(\frac{\mathbf{k}_1 + \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2}t\right)}{\cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega_2 t)} \tilde{\mathbf{S}}_2 \right]\end{aligned}$$

$$= 2 \cos\left(\frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2} t\right) \cdot \cos\left(\frac{\mathbf{k}_1 + \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2} t\right) [\mathcal{S}_1 \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega_1 t) + \mathcal{S}_2 \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega_2 t)] \quad (\text{C82})$$

可见 $\tilde{\mathbf{S}}$ 在方向上与 $\tilde{\mathbf{S}}_1 + \tilde{\mathbf{S}}_2$ 二者之和也不一定同向，也不一定共线，只与

$$\frac{\tilde{\mathbf{S}}_1}{\cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega_1 t)} + \frac{\tilde{\mathbf{S}}_2}{\cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega_2 t)} \quad (\text{C83})$$

共线。同样若简化考虑 $\langle \mathcal{S}_1, \hat{\mathbf{k}}_1 \rangle = \langle \mathcal{S}_2, \hat{\mathbf{k}}_2 \rangle = 0$ ，则此时 $\tilde{\mathbf{S}}$ 与

$$\mathcal{S}_1 \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega_1 t) \hat{\mathbf{k}}_1 + \mathcal{S}_2 \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega_2 t) \hat{\mathbf{k}}_2 \quad (\text{C84})$$

共线，则 $\tilde{\mathbf{S}}$ 与叠加波的相位波的波矢 $\hat{\mathbf{k}}_p$ 也不一定共线；又因 $\hat{\mathbf{k}}_p \cdot \hat{\mathbf{v}}_p = v_p k_p / \omega_p > 0$ ，得此时 $\tilde{\mathbf{S}}$ 与 $\hat{\mathbf{v}}_p$ 也不一定共线。然而同时群速度 $\hat{\mathbf{v}}_g$ 也不一定与 $\hat{\mathbf{v}}_p$ 共线，即 $\hat{\mathbf{v}}_g \cdot \hat{\mathbf{v}}_p$ 同号于 $\Delta\omega \cdot \Delta k$ ，则便可能有 $\tilde{\mathbf{S}}$ 与 $\hat{\mathbf{v}}_g$ 同向。考察

$$\begin{aligned} \tilde{\mathbf{S}} \cdot \mathbf{v}_g &= \left[2 \cos\left(\frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 - \omega_2}{2} t\right) \cdot \cos\left(\frac{\mathbf{k}_1 + \mathbf{k}_2}{2} \cdot \mathbf{r} - \frac{\omega_1 + \omega_2}{2} t\right) \frac{\Delta\omega}{|\Delta\mathbf{k}|} \right] \\ &\times \left[\mathcal{S}_1 \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega_1 t) \hat{\mathbf{k}}_1 \cdot \hat{\Delta\mathbf{k}} + \mathcal{S}_2 \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega_2 t) \hat{\mathbf{k}}_2 \cdot \hat{\Delta\mathbf{k}} \right] \end{aligned} \quad (\text{C85})$$

它正比于

$$H_{01} \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega_1 t) \hat{\mathbf{k}}_1 \cdot (\mathbf{k}_1 - \mathbf{k}_2) + H_{02} \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega_2 t) \hat{\mathbf{k}}_2 \cdot (\mathbf{k}_1 - \mathbf{k}_2) \quad (\text{C86})$$

不是个恒定标量，因此一般而言 $\tilde{\mathbf{S}}, \mathbf{v}_g$ 方向是不同的。可见群速度与能流密度都在一定程度上代表“能量或信息或波包的传递速度”，但二者的方向不一定同向，且在方向和数量关系上均没有必然联系。当然，从令 “ $k_g = 0$ 导致 $v_g \rightarrow \infty$ 而失效，或令 $k_g, \omega_g = 0$ 导致 v_g 取值不定(任意取值)，而此时叠加波的 $\tilde{\mathbf{S}}$ 仍存在”也可知， $\tilde{\mathbf{S}}$ 与 $\hat{\mathbf{v}}_g$ 在大小和方向上均没有必然联系。此时

$$\mathbf{I} = \sum_{i=1}^2 \overline{\tilde{\mathbf{S}}_i} = \sum_{i=1}^2 \left[\mathbf{E}_i^* \times \mathbf{H}_i + (\mathbf{E}_i^* \times \mathbf{H}_i)^* \right] = 2\mathbf{S}_{01} + 2\mathbf{S}_{02} = \frac{1}{2}(\mathcal{S}_1 + \mathcal{S}_2) \quad (\text{C87})$$

附C1.2.8 电各向异性下电矢量无源条件不成立，但电矢量有源不一定无平面波解

可以证明，当 $\tilde{\mathbf{E}}_n$ 的解近似单色平面电磁波时，由于由各阶非线性效应合成的 $\tilde{\mathbf{P}}_n^{NL} = \tilde{\mathbf{D}}_n^{NL}$ 在此前已是单色平面电磁波了，振幅矢量为常矢量，相位部分拥有 $\omega_n = \sum_{i=1}^l \omega_i$ 、 $\mathbf{k}_n = \sum_{i=1}^l \mathbf{k}_i$ 两大电磁波的主要属性，则此时 $\tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 也近似单色平面电磁波，但有两方面的瑕疵，一是波矢中因含有 $\tilde{\mathbf{E}}_n$ 和 $\tilde{\mathbf{P}}_n^{(1)}, \tilde{\mathbf{D}}_n^{(1)}$ 的波矢而不一定为 $\sum_{i=1}^l \mathbf{k}_i$ ，二是

振幅矢量因含有 $\tilde{\mathbf{E}}_n$ 和 $\tilde{\mathbf{P}}_n^{(1)}, \tilde{\mathbf{D}}_n^{(1)}$ 的振幅，而并非完全的常矢量，之后会推出 $\tilde{\mathbf{E}}_n$ 的振幅是随空间缓变的；在完美相位匹配条件下， $\tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 与 $\tilde{\mathbf{E}}_n$ 共用同一对 ω_n, \mathbf{k}_n ^①，但振幅矢量 $\mathbf{P}_{0n}, \mathbf{D}_{0n}$ 大小和方向与 $\tilde{\mathbf{E}}_n$ 的 $\mathbf{E}_{0n} = \mathbf{A}_n$ 不同，且三者的大小均随空间缓变。据无源条件

$$\nabla \cdot \tilde{\mathbf{D}}_n = \tilde{\rho}_f = 0 :$$

$$\begin{aligned} \nabla \cdot \mathbf{D}_n(\mathbf{r}) &= \mathbf{D}_{0n} \cdot \nabla e^{ik_n \cdot \mathbf{r}} + e^{ik_n \cdot \mathbf{r}} \nabla \cdot \mathbf{D}_{0n} \xrightarrow[\nabla \cdot \mathbf{D}_{0n}=0]{\mathbf{D}_{0n} \text{为定矢}} \mathbf{D}_{0n} \cdot \frac{d(e^{ik_n \cdot \mathbf{r}})}{d(k_n \cdot \mathbf{r})} \nabla(k_n \cdot \mathbf{r}) \\ &= \mathbf{D}_{0n} \cdot i e^{ik_n \cdot \mathbf{r}} [\mathbf{k}_n \times (\nabla \times \mathbf{r}) + (\mathbf{k}_n \cdot \nabla) \mathbf{r} + \mathbf{r} \times (\nabla \times \mathbf{k}_n) + (\mathbf{r} \cdot \nabla) \mathbf{k}_n] \\ &\xrightarrow[\nabla \times \mathbf{r} = 0, (\mathbf{k}_n \cdot \nabla) \mathbf{r} = \mathbf{k}_n \cdot \nabla \mathbf{r}]{\mathbf{k}_n \cdot \nabla \mathbf{r} = 0} i[\mathbf{k}_n \cdot \nabla \mathbf{r}] \cdot \mathbf{D}_{0n} e^{ik_n \cdot \mathbf{r}} \xrightarrow[\mathbf{k}_n \cdot \overrightarrow{\mathbf{r}} = \overrightarrow{\mathbf{r}} \cdot \mathbf{k}_n]{\nabla \mathbf{r} = \overrightarrow{\mathbf{r}}} i\mathbf{k}_n \cdot \mathbf{D}_n(\mathbf{r}) \xrightarrow{\nabla \cdot \mathbf{D}_n(\mathbf{r}) = 0} 0 \end{aligned} \quad (C88)$$

上式说明，当 $\tilde{\mathbf{E}}_n$ 的解近似为平面电磁波（包括相位部分为 $e^{i(\mathbf{k}_n \cdot \mathbf{r} - \omega_n t)}$ 且振幅缓变）且满足完美相位匹配时，在无自由电荷源的任何（包括非线性）介质中均普遍满足 $i\mathbf{k}_n \cdot \mathbf{D}_n(\mathbf{r}) \approx \nabla \cdot \mathbf{D}_n(\mathbf{r}) = 0$ 、近似有 $\mathbf{D}_n(\mathbf{r}) \perp \mathbf{k}_n$ ；在特殊坐标系下的各向异性线性介质中

$$\begin{aligned} \tilde{\mathbf{D}}_n &= \tilde{\mathbf{D}}_n^{(1)} = \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \tilde{\mathbf{E}}_n = \epsilon_{nx}^{(1)} \tilde{\mathbf{E}}_{nx} + \epsilon_{ny}^{(1)} \tilde{\mathbf{E}}_{ny} + \epsilon_{nz}^{(1)} \tilde{\mathbf{E}}_{nz} \\ \tilde{\mathbf{E}}_n &= \overset{\rightarrow}{\epsilon}_n^{(1)-1} \cdot \tilde{\mathbf{D}}_n = \epsilon_{nx}^{(1)-1} \tilde{\mathbf{D}}_{nx} + \epsilon_{ny}^{(1)-1} \tilde{\mathbf{D}}_{ny} + \epsilon_{nz}^{(1)-1} \tilde{\mathbf{D}}_{nz} \end{aligned} \quad (C89)$$

以致 $\mathbf{E}_n(\mathbf{r})$ 偏振方向与 $\mathbf{D}_n(\mathbf{r})$ 不同，且大概率有

$$\begin{aligned} \nabla \cdot \mathbf{E}_n(\mathbf{r}) &\approx i\mathbf{k}_n \cdot \mathbf{E}_n(\mathbf{r}) = i[\epsilon_{nx}^{(1)-1} k_{nx} D_{nx}(\mathbf{r}) + \epsilon_{ny}^{(1)-1} k_{ny} D_{ny}(\mathbf{r}) + \epsilon_{nz}^{(1)-1} k_{nz} D_{nz}(\mathbf{r})] \\ &= i\{\epsilon_{nx}^{(1)-1} - \epsilon_{nz}^{(1)-1}\} k_{nx} D_{nx}(\mathbf{r}) + \{\epsilon_{ny}^{(1)-1} - \epsilon_{nz}^{(1)-1}\} k_{ny} D_{ny}(\mathbf{r}) \neq 0 \end{aligned} \quad (C90)$$

即 $\mathbf{E}_n(\mathbf{r}) \not\propto \mathbf{k}_n$ ，除非 $\mathbf{k}_n //$ 为对角化 $\overset{\rightarrow}{\epsilon}_n^{(1)}$ 所选定的三个坐标轴 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 中的任何一个（即 // 介电常数张量椭球的某个主轴），此时 \mathbf{k}_n 只有该方向的分量，同时 $\mathbf{D}_n(\mathbf{r})$ 没有该方向

^① 之前提到过，电场非线性效应(C1)已经决定了各阶 $\tilde{\mathbf{D}}_n^{(I)}$ 连同 $\tilde{\mathbf{D}}_n^{NL}$ 的 ω_n, \mathbf{k}_n ，但 $\tilde{\mathbf{E}}_n$ 的 ω_n, \mathbf{k}_n 并不能与 $\tilde{\mathbf{D}}_n^{NL}$ 的相应量简单地划等号： $\tilde{\mathbf{E}}_n$ 与 $\tilde{\mathbf{D}}_n^{NL}$ 共用 $\tilde{\mathbf{D}}_n^{NL}$ 的 ω_n 是为了将波动方程变为定态；而 $\tilde{\mathbf{E}}_n$ 与 $\tilde{\mathbf{D}}_n^{NL}$ 共用 $\tilde{\mathbf{D}}_n^{NL}$ 的 \mathbf{k}_n 则需要满足完美相位匹配条件。——不过从物理情景的角度，在 $\tilde{\mathbf{D}}_n^{NL}$ 的 $\mathbf{k}_n = \sum_{i=1}^3 k_i$ 由非线性过程决定后（毕竟介质只能在参与混频的入射光场附近，受激并产生响应，因此其波矢由混频光波矢决定是可以理解的），作为本质上也是非线性极化波 $\tilde{\mathbf{P}}_n^{NL}$ 的非线性波源，而其在其光路附近产生的和/差频波 $\tilde{\mathbf{E}}_n$ ，其频率和波矢也应沿用 $\tilde{\mathbf{D}}_n^{NL}$ 的 ω_n, \mathbf{k}_n ，至少其波矢在方向上与 $\tilde{\mathbf{D}}_n^{NL}$ 的 \mathbf{k}_n 同向，且频率与 $\tilde{\mathbf{D}}_n^{NL}$ 的 ω_n 相等，但现实却不完全是这样，在准相位匹配处会发现 $\tilde{\mathbf{E}}_n$ 与 $\tilde{\mathbf{D}}_n^{NL}$ 的波矢方向也可以不同；只有在完美相位匹配条件下，以及 $\tilde{\mathbf{E}}_n, \tilde{\mathbf{P}}_n, \tilde{\mathbf{D}}_n$ 解均是平面电磁波的条件下， $\tilde{\mathbf{E}}_n, \tilde{\mathbf{P}}_n^{NL} = \tilde{\mathbf{D}}_n^{NL}, \tilde{\mathbf{P}}_n^{(1)}, \tilde{\mathbf{D}}_n^{(1)}, \tilde{\mathbf{P}}_n$ ， $\tilde{\mathbf{D}}_n$ 的频率和波矢均相等。

的分量，才有 $\mathbf{k}_n \cdot \mathbf{E}_n(\mathbf{r}) = 0$ ，以至 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) \approx 0$ 。此时对 $\mathbf{D}_n(\mathbf{r})$ 的操作为对其绕 \mathbf{k}_n 旋转并拉伸。

当然，非线性电介质中由于 $\tilde{\mathbf{E}}_n$ 的振幅缓变，即使有 $\mathbf{E}_n(\mathbf{r}) \perp \mathbf{k}_n$ ，也有 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) \approx i\mathbf{k}_n \cdot \mathbf{E}_n(\mathbf{r}) = 0$ ，即即使 $\tilde{\mathbf{E}}_n$ 严格垂直于其波矢，散度也不一定严格为零，反之亦然。不过从 $\tilde{\mathbf{D}}_n = \tilde{\mathbf{D}}_n^{(1)} = \overleftrightarrow{\epsilon}_n^{(1)} \cdot \tilde{\mathbf{E}}_n$ 以及其所导致的波动方程右侧非线性项为零可见，在 $\overrightarrow{\epsilon}_n^{(1)} \rightarrow -1$ 线性电介质中 $\tilde{\mathbf{E}}_n$ 的振幅恒定， $\nabla \cdot \mathbf{E}_n(\mathbf{r})$ 与 $i\mathbf{k}_n \cdot \mathbf{E}_n(\mathbf{r})$ 等价。

附C1.2.8.1 电磁各向同性的非线性电介质中的解收敛到定振幅平面波，电矢量逼近无源，且波矢逼近垂直于电矢量

在电各向同性的线性介质中，由于 $i\mathbf{k}_n \cdot \mathbf{D}_n(\mathbf{r}) = \nabla \cdot \mathbf{D}_n(\mathbf{r}) = 0$ 且 $\mathbf{E}_n(\mathbf{r}) \parallel \mathbf{D}_n(\mathbf{r})$ ，才有严格的 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) = i\mathbf{k}_n \cdot \mathbf{E}_n(\mathbf{r}) = 0$ ，这便圆了式(C58)的大前提，此时电磁各向同性介质中的无源波动方程解才为平面电磁波。此时 $\tilde{\mathbf{D}}_n$ 与 $\tilde{\mathbf{E}}_n$ 的解，均是严格的平面波，不仅振幅不随空间缓变、相位为 $e^{j(\mathbf{k}_n \cdot \mathbf{r} - \omega_n t)}$ ，且严格共用一对 \mathbf{k}_n, ω_n 。但若线性介质是电各向异性的，即使解振幅不随空间缓变，也只有 $i\mathbf{k}_n \cdot \mathbf{D}_n(\mathbf{r}) = \nabla \cdot \mathbf{D}_n(\mathbf{r}) = 0$ ，却因 $\mathbf{E}_n(\mathbf{r}) \not\parallel \mathbf{D}_n(\mathbf{r})$ 而只有 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) = i\mathbf{k}_n \cdot \mathbf{E}_n(\mathbf{r}) \neq 0$ ，并且即使 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) = i\mathbf{k}_n \cdot \mathbf{E}_n(\mathbf{r}) = 0$ ，根据式(C59)，电磁各向异性介质，甚至电各向异性介质中的无源波动方程的解，均无法单独求解，只能对式(C62)令系数行列式为零，且无法作无源近似。

附C1.2.8.2 磁各向同性、电各向异性的非线性电介质中的解的电矢量有源，但其解是否收敛到定振幅平面波与之无关

因此其实在各向异性的线性电介质中，就已经有 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) \neq 0$ 了，非线性电介质中更可能有 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) \neq 0$ ，所以电矢量无源将不会作为求解的前提，因为它太难满足了，毕竟在极化体系中就无法忽略极化电荷的存在；但由于小前提给出波动方程右侧非线性波源项 $\mathbf{P}_n^{NL}(\mathbf{r}) = \mathbf{D}_n^{NL}(\mathbf{r}) = \mathbf{0}$ ，于是非线性介质近似于线性介质，所以至少对 $\nabla \cdot \mathbf{E}_n(\mathbf{r})$ 不用考虑介质的非线性问题；只需将其当做线性介质并考虑其是否各向异性的问题。目前来说，只要非线性电介质的电极化率的线性部分，即二阶介电常数 $\overleftrightarrow{\epsilon}_n^{(1)}$ 是各向同性的（不论高阶介电常数/极化率/非线性系数是否各向同性），即二阶介电常数椭球是球形，则有 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) = 0$ （小前提+一阶各向同性=大前提），且解为平面波；因此上述对波动方程的 $\nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})]$ 部分的化简操作只局限于无双折射现象（非单轴也非双轴）的非线性晶体，否则每个分量方程如(C62)都混有其他分量的物理量成分，每个分量波动方程两两相互关联，电矢量无源时解出来的波是否是无限大平面电磁波，还有待进一步研究；而之后的研究将对此给出肯定的答案：只要介质是磁各向同性的，

那么不论其是否是电各向异性，六大物理量的解均是无限大平面电磁波。并且该结论在满足振幅缓变和完美相位匹配的磁各向同性、电各向异性的非线性光子晶体中也成立。

附C1.2.9 磁各向同性的介质中，即使电各向异性，六大物理量的解仍均为平面波

附C1.2.9.1 \tilde{E}_n 近似定振幅平面波及满足完美相位匹配的两个条件下，电磁均非线性的普遍介质中，波矢形式的微分形式麦氏方程组

类似式(C88)地，据麦氏方程组的第二式，有 $\nabla \cdot \mathbf{B}_n(\mathbf{r}) \approx i\mathbf{k}_n \cdot \mathbf{B}_n(\mathbf{r}) \xrightarrow{\nabla \cdot \mathbf{B}_n(\mathbf{r})=0} 0$ ，则近似有 $\mathbf{B}_n(\mathbf{r}) \perp \mathbf{k}_n$ ，这也是在 \tilde{E}_n 近似为平面电磁波且满足完美相位匹配时，对任何介质均适用的普遍结论；同理 $\nabla \times \mathbf{E}_n(\mathbf{r}) = \nabla e^{ik_n \cdot r} \times \mathbf{E}_{0n} + e^{ik_n \cdot r} \nabla \times \mathbf{E}_{0n} \xrightarrow[\nabla \times \mathbf{E}_{0n}=0]{E_{0n} \text{ 定矢}} \nabla e^{ik_n \cdot r} \times \mathbf{E}_{0n} \xrightarrow{\nabla e^{ik_n \cdot r} = ik_n e^{ik_n \cdot r}} i\mathbf{k}_n \times \mathbf{E}_n(\mathbf{r})$ 以及 $\nabla \times \mathbf{H}_n(\mathbf{r}) \approx i\mathbf{k}_n \times \mathbf{H}_n(\mathbf{r})$ 也在这两个条件满足时成立，于是在这两个条件下，总的有：

$$\left\{ \begin{array}{l} i\mathbf{k}_n \cdot \mathbf{D}_n(\mathbf{r}) \approx \nabla \cdot \mathbf{D}_n(\mathbf{r}) = 0 \\ i\mathbf{k}_n \cdot \mathbf{B}_n(\mathbf{r}) \approx \nabla \cdot \mathbf{B}_n(\mathbf{r}) = 0 \\ i\mathbf{k}_n \times \mathbf{E}_n(\mathbf{r}) e^{-i\omega_n t} \approx \nabla \times \tilde{\mathbf{E}}_n = -\frac{\partial \tilde{\mathbf{B}}_n}{\partial t} = i\omega_n \mathbf{B}_n(\mathbf{r}) e^{-i\omega_n t} \\ i\mathbf{k}_n \times \mathbf{H}_n(\mathbf{r}) e^{-i\omega_n t} \approx \nabla \times \tilde{\mathbf{H}}_n = \frac{\partial \tilde{\mathbf{D}}_n}{\partial t} = -i\omega_n \mathbf{D}_n(\mathbf{r}) e^{-i\omega_n t} \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \mathbf{k}_n \cdot \mathbf{D}_n(\mathbf{r}) \approx 0 \\ \mathbf{k}_n \cdot \mathbf{B}_n(\mathbf{r}) \approx 0 \\ \mathbf{k}_n \times \mathbf{E}_n(\mathbf{r}) \approx \omega_n \mathbf{B}_n(\mathbf{r}) \\ \mathbf{k}_n \times \mathbf{H}_n(\mathbf{r}) \approx -\omega_n \mathbf{D}_n(\mathbf{r}) \end{array} \right. \quad (C91)$$

所有式两边同除以 k_n 并取负（调换物理量与 $\hat{\mathbf{k}}_n$ 的位置）后，将 $\hat{\mathbf{k}}_n \times$ 作用于三式、四式左侧，以三式为例有

$$\mathbf{E}_{n\perp}(\mathbf{r}) := \mathbf{E}_n(\mathbf{r}) - \mathbf{E}_{n\parallel}(\mathbf{r}) = \mathbf{E}_n(\mathbf{r}) - [\hat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] \hat{\mathbf{k}}_n \approx -v_n \hat{\mathbf{k}}_n \times \mathbf{B}_n(\mathbf{r}) = v_n \mathbf{B}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n \quad (C92)$$

则有：

$$\left\{ \begin{array}{l} \mathbf{D}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n \approx 0 \\ \mathbf{B}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n \approx 0 \\ \mathbf{E}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n \approx -v_n \mathbf{B}_n(\mathbf{r}) \\ \mathbf{H}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n \approx v_n \mathbf{D}_n(\mathbf{r}) \end{array} \right. \xrightarrow{\hat{\mathbf{k}}_n \times \text{作用于三式、四式左侧}} \left\{ \begin{array}{l} \mathbf{D}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n \approx 0 \\ \mathbf{B}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n \approx 0 \\ \mathbf{E}_{n\perp}(\mathbf{r}) \approx v_n \mathbf{B}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n \\ \mathbf{H}_{n\perp}(\mathbf{r}) \approx -v_n \mathbf{D}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n \end{array} \right. \quad (C93)$$

附C1.2.9.2 电磁均各向异性的线性介质中，波矢形式麦氏方程组仍无法化简

在(各向异性的)线性介质中，(C92)、(C93)的四个式子，由于振幅不再缓变，且又在解为平面电磁波的条件下，均可将约等号化为等号，并可将 $\mathbf{B}_n(\mathbf{r}) = \overset{\leftrightarrow}{\mu}_n^{(1)} \cdot \overset{\leftrightarrow}{\mathbf{H}}_n(\mathbf{r})$ ， $\mathbf{D}_n(\mathbf{r}) = \overset{\leftrightarrow}{\epsilon}_n^{(1)} \cdot \overset{\leftrightarrow}{\mathbf{E}}_n(\mathbf{r})$ 线性展开；但由于前两个式子因 $\nabla \cdot \mathbf{E}_n(\mathbf{r}) = i\mathbf{k}_n \cdot \mathbf{E}_n(\mathbf{r}) \neq 0$ 、 $\nabla \cdot \mathbf{H}_n(\mathbf{r}) = i\mathbf{k}_n \cdot \mathbf{H}_n(\mathbf{r}) \neq 0$ 而没有进步，所以本身并没有更简洁的表示形式。

附C1.2.9.3 磁各向同性、电各向异性的线性介质中，波矢形式麦氏方程组可化简

进一步地，在磁各向同性、电各向异性的介质中，还额外满足

$\mathbf{B}_n(\mathbf{r}) = \overset{\rightarrow}{\mu}_n^{(1)} \cdot \mathbf{H}_n(\mathbf{r}) = \mu_n^{(1)} \mathbf{H}_n(\mathbf{r})$ ，于是 $\nabla \cdot \mathbf{H}_n(\mathbf{r}) = i \mathbf{k}_n \cdot \mathbf{H}_n(\mathbf{r}) = 0$ 、 $\mathbf{H}_{n\perp}(\mathbf{r}) = \mathbf{H}_n(\mathbf{r})$ ，且(C93)右侧的三式化为 $\mathbf{E}_{n\perp}(\mathbf{r}) = v_n \mu_n^{(1)} \mathbf{H}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n$ ，将(C93)左侧的四式代入其中，并选定合适的坐标系，将介电常数椭球和磁导率椭球同时对角化得 $\mathbf{E}_{n\perp}(\mathbf{r}) = v_n^2 \mu_n^{(1)} \mathbf{D}_n(\mathbf{r}) = \mathbf{D}_n(\mathbf{r}) / \varepsilon_n^{(1)} = \mathbf{D}_n(\mathbf{r}) / \varepsilon_0 \varepsilon_{rn}^{(1)} = \mathbf{D}_n(\mathbf{r}) / \varepsilon_0 n_{ne} |n_{ne}|^{\circledast}$ ，则二、三式可彻底化为：

$$\left\{ \begin{array}{l} \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \mathbf{E}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n = 0 \\ \mathbf{H}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n = 0 \\ \mathbf{E}_{n\perp}(\mathbf{r}) = \frac{1}{\varepsilon_n^{(1)}} \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \mathbf{E}_n(\mathbf{r}) \\ \mathbf{H}_{n\perp}(\mathbf{r}) = -v_n \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \mathbf{E}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \mathbf{D}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n = 0 \\ \mathbf{H}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n = 0 \\ \mathbf{D}_n(\mathbf{r}) = \varepsilon_n^{(1)} \mathbf{E}_{n\perp}(\mathbf{r}) \\ \mathbf{H}_n(\mathbf{r}) = -v_n \mathbf{D}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \mathbf{D}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n = 0 \\ \mathbf{H}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n = 0 \\ \mathbf{D}_n(\mathbf{r}) = \varepsilon_0 n_{ne} |n_{ne}| \mathbf{E}_{n\perp}(\mathbf{r}) \\ \mathbf{H}_n(\mathbf{r}) = -\sqrt{\frac{\varepsilon_n^{(1)}}{\mu_n^{(1)}}} \mathbf{E}_{n\perp}(\mathbf{r}) \times \hat{\mathbf{k}}_n \end{array} \right. \quad (\text{C94})$$

其中 $n_{ne} |n_{ne}| = \varepsilon_{rn}^{(1)} = n_n^2 / \mu_{rn}^{(1)}$ 为光学中常用的折射率，与 $\mu_{rn}^{(1)}$ 、 $\varepsilon_{rn}^{(1)}$ 、 n_n 同正或同负，其中不含相对磁导率（折射率是恒不含真空介电常数和真空磁导率的），与磁无关；但仍适用于磁各向同性的电磁介质。注意其中矢量遇张量先点乘张量，且没有点积交换律 $[\overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \mathbf{E}_n(\mathbf{r})] \cdot \hat{\mathbf{k}}_n \neq \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot [\mathbf{E}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n]$ （后者张量点乘标量甚至是个张量）、 $[\overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \mathbf{E}_n(\mathbf{r})] \times \hat{\mathbf{k}}_n \neq \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot [\mathbf{E}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n]$ ，后者无法交换点积和叉积先后次序的原因：

$$\begin{aligned} & \left[\overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \mathbf{E}_n(\mathbf{r}) \right] \cdot \hat{\mathbf{k}}_n - \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \left[\mathbf{E}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n \right] = \mathbf{E}_n(\mathbf{r}) \times \left\{ \left(\vec{I} \cdot \hat{\mathbf{k}}_n \right) \cdot \left(\begin{bmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{bmatrix} \times (\overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \begin{bmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{bmatrix}) \right) \right\} \\ & -2[\vec{I} \cdot \mathbf{E}_n(\mathbf{r})] \otimes \left\{ \left(\vec{I} \cdot \hat{\mathbf{k}}_n \right) \cdot \left(\begin{bmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{bmatrix} \times (\overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \begin{bmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{bmatrix}) \right) \right\} \end{aligned} \quad (\text{C95})$$

^① 可以有 $\mathbf{E}_{n\perp}(\mathbf{r}) = v_n^2 \mu_n^{(1)} \mathbf{D}_n(\mathbf{r}) = \frac{c^2}{n_n^2} \mu_n^{(1)} \mathbf{D}_n(\mathbf{r})$ ，但不能继续有

$$= \left(\frac{1}{\mu_{nxx}^{(1)} \varepsilon_{nxx}^{(1)}} + \frac{1}{\mu_{nyy}^{(1)} \varepsilon_{nyy}^{(1)}} + \frac{1}{\mu_{nzz}^{(1)} \varepsilon_{nzz}^{(1)}} \right) \mu_n^{(1)} \mathbf{D}_n(\mathbf{r}) = \left(\frac{1}{\varepsilon_{nxx}^{(1)}} + \frac{1}{\varepsilon_{nyy}^{(1)}} + \frac{1}{\varepsilon_{nzz}^{(1)}} \right) \mathbf{D}_n(\mathbf{r}) = \frac{1}{\varepsilon_0 \varepsilon_{rn}^{(1)}} \mathbf{D}_n(\mathbf{r}) =$$

$\frac{1}{\varepsilon_0 n_{ne} |n_{ne}|} \mathbf{D}_n(\mathbf{r})$ ，因为 $v_n^2 = \frac{c^2}{n_n^2}$ 是成立的，但 $\frac{1}{n_n^2} = \frac{1}{n_{nex}^2} + \frac{1}{n_{ney}^2} + \frac{1}{n_{nez}^2}$ 不成立；之后才可导

出 $\varepsilon_{rn}^{(1)} = n_{ne}^2$ 的表达式。

对于电各向异性的介质不等于零，只有当介质同时是电各向同性时，才有：

$$\begin{aligned} & \left| \begin{array}{ccc} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \varepsilon_{nx}^{(1)} E_{nx} & \varepsilon_{ny}^{(1)} E_{ny} & \varepsilon_{nz}^{(1)} E_{nz} \\ \frac{k_{nx}}{k_n} & \frac{k_{ny}}{k_n} & \frac{k_{nz}}{k_n} \end{array} \right| - \left| \begin{array}{ccc} \varepsilon_{nx}^{(1)} \mathbf{i} & \varepsilon_{ny}^{(1)} \mathbf{j} & \varepsilon_{nz}^{(1)} \mathbf{k} \\ E_{nx} & E_{ny} & E_{nz} \\ \frac{k_{nx}}{k_n} & \frac{k_{ny}}{k_n} & \frac{k_{nz}}{k_n} \end{array} \right| = \left| \begin{array}{ccc} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ E_{nx} & E_{ny} & E_{nz} \\ \frac{1}{\varepsilon_{ny}^{(1)}} & \frac{1}{\varepsilon_{nz}^{(1)}} & \frac{1}{\varepsilon_{nx}^{(1)}} \end{array} \right| \left| \begin{array}{c} k_{nx} \\ k_{ny} \\ k_{nz} \end{array} \right| \\ & - 2 \left| \begin{array}{ccc} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ E_{nx} \mathbf{jj} & E_{ny} \mathbf{kk} & E_{nz} \mathbf{ii} \\ \frac{1}{\varepsilon_{ny}^{(1)}} & \frac{1}{\varepsilon_{nz}^{(1)}} & \frac{1}{\varepsilon_{nx}^{(1)}} \end{array} \right| = \mathbf{0} \end{aligned} \quad (C96)$$

因此电各向异性的最左侧四式无法像三式那样化到最简，只能将二、三式代入其中以略微简化。

附C1.2.9.4 磁各向同性线性电磁介质中，六大物理量均为严格的定振幅平面波，即使电矢量有源

现便可以证明在磁各向同性的线性介质中，无边界条件时，六大物理量的解均是严格的平面电磁波：在“ $\mathbf{E}_n(\mathbf{r})$ 为定振幅的平面电磁波的复振幅”大条件下，利用“ ∇ 作用于定矢得零，且 ∇ 会作用于任何含 \mathbf{r} 的函数，但 \mathbf{k}_n 只作用于矢量”，有：

$$\begin{aligned} \nabla \times [\nabla \times \mathbf{E}_n(\mathbf{r})] &= \nabla \times [\mathbf{i} \mathbf{k}_n \times \mathbf{E}_n(\mathbf{r})] = \mathbf{i} \nabla \times [(\mathbf{k}_n \times \mathbf{E}_{0n}) e^{i\mathbf{k}_n \cdot \mathbf{r}}] = \mathbf{i} \nabla e^{i\mathbf{k}_n \cdot \mathbf{r}} \times [\mathbf{k}_n \times \mathbf{E}_{0n}] \\ &= \mathbf{i} \mathbf{i} \mathbf{k}_n \times [\mathbf{k}_n \times \mathbf{E}_{0n}] = \mathbf{k}_n \times [\mathbf{E}_{0n} \times \mathbf{k}_n] = k_n^2 \mathbf{E}_{n\perp}(\mathbf{r}) = k_n^2 \frac{1}{\varepsilon_n^{(1)}} \mathbf{D}_n(\mathbf{r}) \end{aligned} \quad (C97)$$

将其与磁各向同性条件 $\vec{\mu}_n^{(1)}(\omega_n) = \mu_n^{(1)} \vec{I}$ 一并代入式(C52)得 $k_n^2 \mathbf{D}_n(\mathbf{r}) / \varepsilon_n^{(1)} - \omega_n^2 \mu_n^{(1)} \mathbf{D}_n(\mathbf{r}) = \omega_n^2 \mu_n^{(1)} \mathbf{P}_n^{NL}(\mathbf{r}) = \mathbf{0}$ ，得到 $k_n^2 / \varepsilon_n^{(1)} = \omega_n^2 \mu_n^{(1)}$ ，即 $1 / \mu_n^{(1)} \varepsilon_n^{(1)} = \omega_n^2 / k_n^2 = v_n^2$ ，这既是各参量遵循的正确关系式，又是磁各向同性线性介质中的电场波动方程（只不过其连同这里的麦氏方程组都是“波矢形式”）。该结果说明此时“定振幅的 $\mathbf{E}_n(\mathbf{r}) = \mathbf{E}_{0n} e^{i\mathbf{k}_n \cdot \mathbf{r}}$ 确实是波动方程的一个解”，该结论的成立不依赖于 $\mathbf{E}_n(\mathbf{r})$ 是否无源，即不需大条件成立；只依赖于无非线性波源，即小条件成立即可。

而当 $\mathbf{E}_n(\mathbf{r})$ 成为平面电磁波后，据 $\mathbf{E}_n(\mathbf{r}), \mathbf{P}_n(\mathbf{r}), \mathbf{D}_n(\mathbf{r})$ 间的线性关系，此时 $\mathbf{P}_n(\mathbf{r}), \mathbf{D}_n(\mathbf{r})$ 也为平面电磁波；再据式(C94)的式四，得此时 $\mathbf{H}_n(\mathbf{r}) = -\nu_n \mathbf{D}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n$ 也是平面电磁波，再据 $\mathbf{H}_n(\mathbf{r}), \mathbf{M}_n(\mathbf{r}), \mathbf{B}_n(\mathbf{r})$ 间的线性关系，此时 $\mathbf{M}_n(\mathbf{r}), \mathbf{B}_n(\mathbf{r})$ 也为平面电磁波。

附C1.2.9.5 电各向异性磁各向同性线性介质中的菲涅尔波法线方程及其导出的两个折射率曲面

(1) 菲涅尔波法线方程

将 $\varepsilon_n^{(1)} > 0$ 条件下的三式 $D_n(\mathbf{r}) = \varepsilon_0 n_{ne}^2 \mathbf{E}_{n\perp}(\mathbf{r}) = \varepsilon_0 n_{ne}^2 \left\{ \mathbf{E}_n(\mathbf{r}) - [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] \widehat{\mathbf{k}}_n \right\}$ 写成分量形式，其中之一为

$$\begin{aligned} D_{nx} &= \varepsilon_0 n_{ne}^2 \left\{ E_{nx} - [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] k_{n0x} \right\} = \varepsilon_0 n_{ne}^2 \left\{ \frac{D_{nx}}{\varepsilon_{nx}^{(1)}} - [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] k_{n0x} \right\} \\ &= n_{ne}^2 \left\{ \frac{D_{nx}}{\varepsilon_{nx}^{(1)}} - \varepsilon_0 [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] k_{n0x} \right\} \end{aligned} \quad (C98)$$

移项合并同类项得：

$$D_{nx} \left(\frac{1}{n_{ne}^2} - \frac{1}{\varepsilon_{nx}^{(1)}} \right) = D_{nx} \left(\frac{1}{n_{ne}^2} - \frac{1}{n_{nex}^2} \right) = -\varepsilon_0 [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] k_{n0x} \quad (C99)$$

则

$$\begin{aligned} D_{nx} &= -\varepsilon_0 [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] \cdot \frac{k_{n0x}}{\frac{1}{n_{ne}^2} - \frac{1}{n_{nex}^2}} = -\varepsilon_0 [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] \frac{1}{\mu_{rn}^{(1)}} \frac{k_{n0x}}{\frac{1}{n_n^2} - \frac{1}{n_{nx}^2}} \\ &\neq -\varepsilon_0 [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] \frac{c^2}{\mu_{rn}^{(1)} v_n^2 - v_{nx}^2} \neq -\varepsilon_0 [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] \frac{c^2}{\mu_{rn}^{(1)} \omega^2} \frac{k_{n0x}}{\frac{1}{k_n^2} - \frac{1}{k_{nx}^2}} \quad (C100) \end{aligned}$$

同理可得 D_{ny}, D_{nz} ，将三者代入(C94)的一式，有：

$$\mathbf{D}_n(\mathbf{r}) \cdot \widehat{\mathbf{k}}_n = D_{nx} k_{n0x} + D_{ny} k_{n0y} + D_{nz} k_{n0z}$$

^① 注意后两式不成立，因为 $v_{nx}^2 = \frac{1}{\mu_{nx}^{(1)} \varepsilon_{nx}^{(1)}}$ 、 $v_{nx}^2 = \frac{\omega^2}{k_{nx}^2}$ 等均不成立，否则将会有

$\frac{1}{n_n^2} = \frac{1}{n_{nx}^2} + \frac{1}{n_{ny}^2} + \frac{1}{n_{nz}^2}$ 成立，这与折射率椭球和曲面矛盾；也将有 $\frac{1}{k_n^2} = \frac{1}{k_{nx}^2} + \frac{1}{k_{ny}^2} + \frac{1}{k_{nz}^2}$ 成立，

与 $k_n^2 = k_{nx}^2 + k_{ny}^2 + k_{nz}^2$ 矛盾；因此最多将后两式分别写作 $= -\varepsilon_0 [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] \frac{c^2}{\mu_{rn}^{(1)} v_n^2 - \frac{c^2}{n_{nx}^2}}$ ，

$= -\varepsilon_0 [\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})] \frac{c^2}{\mu_{rn}^{(1)} \omega^2} \frac{k_{n0x}}{\frac{1}{k_n^2} - \frac{c^2}{\omega^2 n_{nx}^2}}$ 也就是 $v_{nx}^2, \frac{1}{k_{nx}^2}$ 最终必须回归到“其分别所代表的” $\frac{c^2}{n_{nx}^2}$ ，

$\frac{c^2}{\omega^2 n_{nx}^2}$ ，而 $v_n^2, \frac{1}{k_n^2}$ 可以保留；在这里出现它们，是为了说明通过折射率曲面也可得出相速

(下页续)

$$-\varepsilon_0 \left[\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r}) \right] \left[\frac{k_{n0x}^2}{\frac{1}{n_{ne}^2} - \frac{1}{n_{nex}^2}} + \frac{k_{n0y}^2}{\frac{1}{n_{ne}^2} - \frac{1}{n_{ney}^2}} + \frac{k_{n0z}^2}{\frac{1}{n_{ne}^2} - \frac{1}{n_{nez}^2}} \right] = 0 \quad (\text{C101})$$

其中一般有 $\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r}) \neq 0$ ，即得菲涅尔波法线方程：

$$\frac{k_{n0x}^2}{\frac{1}{n_{ne}^2} - \frac{1}{n_{nex}^2}} + \frac{k_{n0y}^2}{\frac{1}{n_{ne}^2} - \frac{1}{n_{ney}^2}} + \frac{k_{n0z}^2}{\frac{1}{n_{ne}^2} - \frac{1}{n_{nez}^2}} = 0 \quad (\text{C102})$$

其中 $k_{n0x} := |\widehat{\mathbf{k}}_{n_x}| = k_{nx}/k_n$ 。注意该方程既非折射率椭球(方程)，也非折射率曲面。它仅暗示折射率、相速度、波矢，三者的大小，或许均与光波波矢方向有关，因而该方程的解又叫（光学意义上的）折射率曲面 $((\widehat{\mathbf{k}}_n, n_{ne})$ 曲面），并可导出相速度曲面 $((\widehat{\mathbf{k}}_n, v_n)$ 曲面）、波矢(曲)面 $((\widehat{\mathbf{k}}_n, k_n)$ 曲面）；之后会看到，该方程的解有两个，因此折射率曲面也有两个。

(2) 第一种方法：通过菲涅尔波法线方程得出 o 光 e 光的折射率曲面

将菲涅尔波法线方程展开得：

$$\begin{aligned} & [n_{nex}^2 (n_{ne} k_{n0x})^2 + n_{ney}^2 (n_{ne} k_{n0y})^2 + n_{nez}^2 (n_{ne} k_{n0z})^2] \cdot [(n_{ne} k_{n0x})^2 + (n_{ne} k_{n0y})^2 + (n_{ne} k_{n0z})^2] \\ & - [n_{nex}^2 (n_{ney}^2 + n_{nez}^2) (n_{ne} k_{n0x})^2 + n_{ney}^2 (n_{nez}^2 + n_{nex}^2) (n_{ne} k_{n0y})^2 \\ & + n_{nez}^2 (n_{nex}^2 + n_{ney}^2) (n_{ne} k_{n0z})^2] + n_{nex}^2 n_{ney}^2 n_{nez}^2 = 0 \end{aligned} \quad (\text{C103})$$

因式分解得：

$$\begin{aligned} & [(n_{ne} k_{n0x})^2 + (n_{ne} k_{n0y})^2 + (n_{ne} k_{n0z})^2 - n_{nex}^2] \cdot [n_{nex}^2 (n_{ne} k_{n0x})^2 + n_{ney}^2 (n_{ne} k_{n0y})^2 + n_{nez}^2 (n_{ne} k_{n0z})^2 - n_{ney}^2 n_{nez}^2] \\ & + [n_{nex}^2 - n_{ney}^2] \cdot [n_{nex}^2 - n_{nez}^2] \cdot (n_{ne} k_{n0x})^2 = 0 \end{aligned} \quad (\text{C104})$$

设单轴晶体的光轴沿 z 方向，则此时无论三维直角坐标系的 x, y 轴沿什么方向，晶体的折射率椭球均已主轴化： $n_{nex} = n_{ney} =: n_{ne\perp}$ 、 $n_{nez} =: n_{ne//}$ ，代入得：

$$\begin{aligned} & [(n_{ne} k_{n0x})^2 + (n_{ne} k_{n0y})^2 + (n_{ne} k_{n0z})^2 - n_{ne\perp}^2] \\ & \cdot [n_{ne\perp}^2 (n_{ne} k_{n0x})^2 + n_{ne\perp}^2 (n_{ne} k_{n0y})^2 + n_{ne//}^2 (n_{ne} k_{n0z})^2 - n_{ne\perp}^2 n_{ne//}^2] = 0 \end{aligned} \quad (\text{C105})$$

方程有两个解，每一个解均对应一个折射率曲面：

$$(n_{ne} k_{n0x})^2 + (n_{ne} k_{n0y})^2 + (n_{ne} k_{n0z})^2 = n_{ne\perp}^2$$

度曲面与波矢曲面，但注意其方程写法。

$$\frac{(n_{ne0} k_{ne0x})^2 + (n_{ne0} k_{ne0y})^2}{n_{ne\parallel}^2} + \frac{(n_{ne0} k_{ne0z})^2}{n_{ne\perp}^2} = 1 \quad (C106)$$

考察 $\perp x$ 轴入射晶体的光波（晶体或坐标系均可绕 z 轴 or 光轴旋转，以使得任何波矢量总 \perp 其旋转后的 x 轴；此时 $\hat{\mathbf{k}}_n$ 与 \mathbf{k} 的公垂线// \mathbf{i} ）： $k_{nx} = 0, k_{ny} = \sin \theta_n, k_{nz} = \cos \theta_n$ ，其中 $\theta_n = \angle \hat{\mathbf{k}}_n, \mathbf{k}$ ，则两个解变为：

$$n_{neo}^2 = n_{ne\perp}^2 \text{, } \frac{1}{n_{ne\parallel}^2} = \frac{\sin^2 \theta_n}{n_{ne\parallel}^2} + \frac{\cos^2 \theta_n}{n_{ne\perp}^2} \quad (C107)$$

由于 $\varepsilon_n^{(1)} > 0$ ，则还有

$$n_{neo} = n_{ne\perp} \text{, } n_{ne\parallel}(\theta_n) = \frac{n_{ne\parallel} n_{ne\perp}}{\sqrt{n_{ne\perp}^2 \sin^2 \theta_n + n_{ne\parallel}^2 \cos^2 \theta_n}} \quad (C108)$$

其中 $n_{ne\parallel}(\pi/2) = n_{ne\parallel}$ 、 $n_{ne\parallel}(0) = n_{ne\parallel}(\pi) = n_{ne\perp} = n_{neo}$ 。

注意式(C108)这两个解是“或”而非“与”的关系，但既然二者都“可能”存在，现实就是二者“必定”同时存在，即最终导致单轴晶体的双折射现象；作为折射率曲面，二者不仅在数量上与折射率椭球不同，且折射率曲面只能考察波矢或相速度方向上的等效标量折射率，不能考察麦氏方程组中六大物理量的偏振方向上的等效标量折射率，因此比折射率椭球稍微低一个层次；但它的数学形式与折射率椭球完全一致，这方面也印证了二者同根同源，且都没有问题。

(3) 第二种方法：通过电矢量法向极化方程解出 o 光 e 光的折射率曲面，并同时得到两偏振光的其他性质

还有第二种方法得到相应的结论：从(C94)左侧的三式有

$$\mathbf{E}_n(\mathbf{r}) - \left[\hat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r}) \right] \hat{\mathbf{k}}_n = \frac{1}{\varepsilon_n^{(1)}} \vec{\mathbf{e}}_n^{(1)} \cdot \mathbf{E}_n(\mathbf{r}) \quad (C109)$$

根据(C94)左侧的一式有 $\hat{\mathbf{k}}_n \cdot \vec{\mathbf{e}}_n^{(1)} \cdot \mathbf{E}_n(\mathbf{r}) = k_{n0x} \varepsilon_{nx}^{(1)} E_{nx} + k_{n0y} \varepsilon_{ny}^{(1)} E_{ny} + k_{n0z} \varepsilon_{nz}^{(1)} E_{nz} = 0$ ，得

$$\hat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r}) = k_{n0x} E_{nx} + k_{n0y} E_{ny} + k_{n0z} E_{nz} = \left(1 - \frac{\varepsilon_{ny}^{(1)}}{\varepsilon_{nx}^{(1)}} \right) k_{n0y} E_{ny} + \left(1 - \frac{\varepsilon_{nz}^{(1)}}{\varepsilon_{nx}^{(1)}} \right) k_{n0z} E_{nz} \quad (C110)$$

代入其中有：

$$\begin{bmatrix} \left(1 - \frac{\varepsilon_{nx}^{(1)}}{\varepsilon_n^{(1)}}\right) \mathbf{E}_{nx} \\ \left(1 - \frac{\varepsilon_{ny}^{(1)}}{\varepsilon_n^{(1)}}\right) \mathbf{E}_{ny} \\ \left(1 - \frac{\varepsilon_{nz}^{(1)}}{\varepsilon_n^{(1)}}\right) \mathbf{E}_{nz} \end{bmatrix} - \left[\left(1 - \frac{\varepsilon_{ny}^{(1)}}{\varepsilon_{nx}^{(1)}}\right) k_{n0y} E_{ny} + \left(1 - \frac{\varepsilon_{nz}^{(1)}}{\varepsilon_{nx}^{(1)}}\right) k_{n0z} E_{nz} \right] \begin{bmatrix} \mathbf{k}_{n0x} \\ \mathbf{k}_{n0y} \\ \mathbf{k}_{n0z} \end{bmatrix} = \mathbf{0} \quad (\text{C111})$$

将三个分量方程整理成关于 $\mathbf{E}_n(\mathbf{r})$ 的三分量的方程，有：

$$\begin{bmatrix} 1 - \frac{\varepsilon_{nx}^{(1)}}{\varepsilon_n^{(1)}} & -\left(1 - \frac{\varepsilon_{ny}^{(1)}}{\varepsilon_{nx}^{(1)}}\right) k_{n0x} k_{n0y} & -\left(1 - \frac{\varepsilon_{nz}^{(1)}}{\varepsilon_{nx}^{(1)}}\right) k_{n0x} k_{n0z} \\ 0 & \left(1 - \frac{\varepsilon_{ny}^{(1)}}{\varepsilon_n^{(1)}}\right) - \left(1 - \frac{\varepsilon_{ny}^{(1)}}{\varepsilon_{nx}^{(1)}}\right) k_{n0y} k_{n0y} & -\left(1 - \frac{\varepsilon_{nz}^{(1)}}{\varepsilon_{nx}^{(1)}}\right) k_{n0y} k_{n0z} \\ 0 & -\left(1 - \frac{\varepsilon_{ny}^{(1)}}{\varepsilon_{nx}^{(1)}}\right) k_{n0z} k_{n0y} & \left(1 - \frac{\varepsilon_{nz}^{(1)}}{\varepsilon_n^{(1)}}\right) - \left(1 - \frac{\varepsilon_{nz}^{(1)}}{\varepsilon_{nx}^{(1)}}\right) k_{n0z} k_{n0z} \end{bmatrix} \begin{bmatrix} E_{nx} \\ E_{ny} \\ E_{nz} \end{bmatrix} = \mathbf{0} \quad (\text{C112})$$

要想该方程有非零解，系数矩阵的行列式的值为零。对于单轴晶体而言，进一步简化为：

$$\begin{bmatrix} 1 - \frac{\varepsilon_{n\perp}^{(1)}}{\varepsilon_n^{(1)}} & 0 & -\left(1 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}}\right) k_{n0x} k_{n0z} \\ 0 & 1 - \frac{\varepsilon_{n\perp}^{(1)}}{\varepsilon_n^{(1)}} & -\left(1 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}}\right) k_{n0y} k_{n0z} \\ 0 & 0 & k_{n0x}^2 + k_{n0y}^2 + \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}} k_{n0z}^2 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_n^{(1)}} \end{bmatrix} \begin{bmatrix} E_{nx} \\ E_{ny} \\ E_{nz} \end{bmatrix} = \mathbf{0} \quad (\text{C113})$$

其中

$$\frac{\varepsilon_{n\perp}^{(1)}}{\varepsilon_n^{(1)}} = \frac{\omega^2 \mu_n^{(1)} \varepsilon_{n\perp}^{(1)}}{k^2}, \quad \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_n^{(1)}} = \frac{\omega^2 \mu_n^{(1)} \varepsilon_{n//}^{(1)}}{k^2}$$

$$k_{n0x}^2 + k_{n0y}^2 + \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}} k_{n0z}^2 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_n^{(1)}} = \left(1 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_n^{(1)}}\right) - \left(1 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}}\right) k_{n0z} k_{n0z} \quad (\text{C114})$$

且有

$$\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r}) = k_{n0x} E_{nx} + k_{n0y} E_{ny} + k_{n0z} E_{nz} = \left(1 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}}\right) k_{n0z} E_{nz}$$

$$\frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}} = 1 - \frac{\widehat{\mathbf{k}}_n \cdot \mathbf{E}_n(\mathbf{r})}{k_{n0z} E_{nz}} \quad (\text{C115})$$

令其系数矩阵行列式为零，得两个解：

$$\varepsilon_{no}^{(1)} = \varepsilon_{n\perp}^{(1)} \quad \text{或} \quad k_{no}^2 = \omega^2 \mu_n^{(1)} \varepsilon_{n\perp}^{(1)}$$

$$\frac{\varepsilon_{ne}^{(1)}(k_{ne0x}^2 + k_{ne0y}^2)}{\varepsilon_{n//}^{(1)}} + \frac{\varepsilon_{ne}^{(1)} k_{ne0z}^2}{\varepsilon_{n\perp}^{(1)}} = 1 \quad \text{或} \quad k_{nex}^2 + k_{ney}^2 + \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}} k_{nez}^2 = \omega^2 \mu_n^{(1)} \varepsilon_{n//}^{(1)} \quad (\text{C116})$$

设 $\theta_{no} = \langle \hat{k}_{no}, \mathbf{k} \rangle$ 、 $\theta_{ne} = \langle \hat{k}_{ne}, \mathbf{k} \rangle$ ，则

$$\varepsilon_{nro}^{(1)} = \varepsilon_{nr\perp}^{(1)} \quad \text{且} \quad \frac{1}{\varepsilon_{ne}^{(1)}} = \frac{\sin^2 \theta_{ne}}{\varepsilon_{n//}^{(1)}} + \frac{\cos^2 \theta_{ne}}{\varepsilon_{n\perp}^{(1)}} \quad \text{即} \quad \frac{1}{\varepsilon_{nre}^{(1)}} = \frac{\sin^2 \theta_{ne}}{\varepsilon_{nr//}^{(1)}} + \frac{\cos^2 \theta_{ne}}{\varepsilon_{nr\perp}^{(1)}} \quad (\text{C117})$$

结果是一样的。此外，e 光的折射率曲面可写作：

$$\left(1 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{ne}^{(1)}}\right) - \left(1 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}}\right) k_{ne0z} k_{ne0z} = 0 \quad (\text{C118})$$

于是还有

$$1 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}} = \frac{1}{k_{ne0z} k_{ne0z}} \left(1 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{ne}^{(1)}}\right)$$

$$\hat{k}_{ne} \cdot \mathbf{E}_{ne}(\mathbf{r}) = \left(1 - \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{ne}^{(1)}}\right) \frac{E_{nez}}{k_{ne0z}} \quad (\text{C119})$$

①. o 光的电矢量方向垂直于 o 光主平面；e 光的电矢量方向平行于 e 光主平面，只有当 o 光 e 光主平面重合时，二者的电矢量才相互垂直

对于寻常波，将 $\varepsilon_{no}^{(1)} = \varepsilon_{n\perp}^{(1)}$ 代入式(C112)，得 $E_{noz} = 0$ ，可见寻常波的偏振方向 $\mathbf{E}_{no}(\mathbf{r}) \perp \mathbf{k}$ ；再将其代入 $\mathbf{D}_{no}(\mathbf{r}) = \overset{\rightarrow}{\varepsilon}_{no}^{(1)} \cdot \mathbf{E}_{no}(\mathbf{r})$ 和(C94)的式一，介电常数张量退化为标量，得 $\mathbf{E}_{no}(\mathbf{r}) // \mathbf{D}_{no}(\mathbf{r}) \perp \hat{k}_{no}$ ，可知此时 $\mathbf{E}_{no}(\mathbf{r}), \mathbf{D}_{no}(\mathbf{r}) \perp$ 其 \mathbf{k}, \hat{k}_{no} 所构成的平面，即 o 光的主平面；若定义 \mathbf{k}, \hat{k}_{no} 所构成的平面为 $y_o - o - z$ 面，则此时 $\mathbf{E}_{no}(\mathbf{r}) = \mathbf{E}_{nox}(\mathbf{r}) // \mathbf{i}_{no}$ 、 $E_{noy} = 0$ ，且 $k_{nox} = 0$ 。

对于非寻常波，将其折射率曲面式(C116)与 $k_{ne0x} = 0$ 代入式(C113)，查看矩阵元第一行所代表的方程，得 $E_{nex} = 0$ ，说明 $\mathbf{E}_{ne}(\mathbf{r}) \perp \mathbf{i}_{ne}$ ，且由于 $\mathbf{E}_{ne}(\mathbf{r}), \mathbf{D}_{ne}(\mathbf{r})$ 共面 ((C94)的式三)，此时 $\mathbf{E}_{ne}(\mathbf{r}), \mathbf{D}_{ne}(\mathbf{r})$ 均在 \mathbf{k}, \hat{k}_{ne} 所构成的平面，即 e 光的主平面内。

当入射面与晶体的主截面（晶体光轴与晶面的法线确定的平面）重合时，o 光与 e 光的主平面即 $y_o - o - z$ 面与 $y_e - o - z$ 面重合且与主截面重合，此时才有 $\mathbf{i}_{no} // \mathbf{i}_{ne}$ 、 $\mathbf{j}_{no} // \mathbf{j}_{ne}$ ，二者同时同向平行(或反向平行)，导致无论在 o 光的 $o - x_o y_o z$ 坐标系下，还是在 $o - x_e y_e z$ 系下，均有 $k_{nox} = k_{nex} = 0$ 且 $\mathbf{E}_{no}(\mathbf{r}), \mathbf{D}_{no}(\mathbf{r}) \perp \mathbf{E}_{ne}(\mathbf{r}), \mathbf{D}_{ne}(\mathbf{r})$ 。

但尽管 o 光和 e 光的主平面重合了，在斜入射的情况下二者的波矢在晶体内仍然

不平行，即 $\widehat{\mathbf{k}_{no}} \nparallel \widehat{\mathbf{k}_{ne}}$ 、 $\theta_{no} \neq \theta_{ne}$ ，二者的夹角需要用斯涅耳作图法作出波矢曲面，并利用折射定律确定两束折射光的波法线方向后具体分析。斯涅耳作图法配合折射、反射定律也可确定两束晶体内部可能存在的反射光的波矢大小和相速度大小，以及晶体内部的两束可能存在的反射光的波矢方向、波矢大小，以及相速度大小。至于两束可能存在的折射或反射光是否存在，要看入射光射入晶体后，在 o 光和 e 光振动方向上是否均有相应的分量。

②. 第二种方法更完备，但两种方法所得的两个折射率曲面均只能给出波矢方向上的波矢大小，不可给出极化方向上的极化程度，极化大小需由折射率椭球决定

从中可见，对于 o 光，由于 $\mathbf{E}_{no}(\mathbf{r}) \perp \widehat{\mathbf{k}_{no}}$ ，导致菲涅尔波法线方程的第一种推导方法所得到的前提 $\widehat{\mathbf{k}_n} \cdot \mathbf{E}_n(\mathbf{r}) \neq 0$ 不成立，以至于第一种方法没法得到菲涅尔波法线方程的 o 光折射率曲面和波矢曲面，只能得到对于 o 光有 $\mathbf{E}_{no}(\mathbf{r}) \perp \widehat{\mathbf{k}_{no}}$ ；但第二种方法就绕过了这个问题，并直接得到 $\varepsilon_{no}^{(1)} = \varepsilon_{n\perp}^{(1)}$ ，但该折射率曲面仍然只适用于波矢和相速度矢量方向的折射率，并只能用于确定波矢和相速度的方向和大小，不可用于确定 o 光的极化方向和大小，它需要由折射率椭球确定：将 $D_{noz} = \varepsilon_{nz}^{(1)} E_{noz} = 0$ 代入

$$\begin{aligned} \frac{(n_{ne}x_0)^2}{n_{nex}^2} + \frac{(n_{ne}y_0)^2}{n_{ney}^2} + \frac{(n_{ne}z_0)^2}{n_{nez}^2} &= \frac{(n_{neo}D_{nox}/D_{no})^2}{n_{ne\perp}^2} + \frac{(n_{neo}D_{noy}/D_{no})^2}{n_{ne\perp}^2} + \frac{(n_{neo}D_{noz}/D_{no})^2}{n_{ne\parallel}^2} \\ &= \frac{n_{neo}^2(D_{nox}^2 + D_{noy}^2)/D_{no}^2}{n_{ne\perp}^2} = 1 \end{aligned} \quad (C120)$$

即有 $n_{neo} = n_{ne\perp}$ ；而若考察 $(x_0, y_0, z_0) = (E_{nox}/E_{no}, E_{noy}/E_{no}, E_{noz}/E_{no})$ ，也有相同的结论，这说明由于 o 光的偏振方向 $\mathbf{E}_{no}(\mathbf{r})$ 恒 $\perp \mathbf{k}$ （而不是 $\widehat{\mathbf{k}_{no}}$ ），而该方向上的等效标量折射率为恒定值 $n_{ne\perp}$ ，也就解释了极化为该方向的 $\mathbf{D}_{no}(\mathbf{r}) \perp \mathbf{k}$ ；在这样的“绕 \mathbf{k} 轴各向同性”环境下，折射率和介电常数不变，o 光的波动方程退化为电磁均各向同性的线性介质中的波动方程，而该方程在无源条件下的解是无源的，则可确定激发此极化的原偏振方向 $\mathbf{E}_{no}(\mathbf{r})$ 无源且 $\perp \widehat{\mathbf{k}_{no}}$ ；或将 $\varepsilon_{no}^{(1)} = \varepsilon_{n\perp}^{(1)}$ 代入 $\mathbf{E}_{no}(\mathbf{r})$ 到 $\mathbf{D}_{no}(\mathbf{r})$ 的极化式得

$\mathbf{E}_{no}(\mathbf{r}) \parallel \mathbf{D}_{no}(\mathbf{r})$ ，而又因 $\mathbf{D}_{no}(\mathbf{r}) \perp \widehat{\mathbf{k}_{no}}$ ，也恒有 $\mathbf{E}_{no}(\mathbf{r}) \perp \widehat{\mathbf{k}_{no}}$ ，这一方面印证了折射率椭球所确定的折射率确实可用于确定极化方向和大小，并且上一段和这里得出结论

$\mathbf{E}_{no}(\mathbf{r}) \parallel \mathbf{D}_{no}(\mathbf{r})$ 所需代入的 $\varepsilon_{no}^{(1)} = \varepsilon_{n\perp}^{(1)}$ ，本质上来自于折射率椭球而不是折射率曲面；另一方面也印证了这里的 o 光解在偏振方向上确实和平常光一样，并且因此

$\tilde{\mathbf{S}}_{no} = \tilde{\mathbf{E}}_{no} \times \tilde{\mathbf{H}}_{no}$ 将同向于 $\widehat{\mathbf{k}_{no}}$ 。

③. 考虑 o 光的(波矢形式)波动方程，可知 o 光的折射率曲面本质上来源于折射

率椭球中 o 光独特的偏振/极化方向上的极化程度

但上述考虑中存在一个看似细微却后果严重的问题：o 光是否是波动方程的解？若菲涅尔波法线方程、折射率椭球均允许 o 光的存在，但波动方程不允许，那也没法存在 o 光：现已知磁各向同性线性介质中，波动方程最终可化为 $1/\mu_n^{(1)}\epsilon_n^{(1)} = \omega_n^2/k_n^2 = v_n^2$ ，与 $\mathbf{D}_n(\mathbf{r}) = \vec{\epsilon}_n^{(1)} \cdot \vec{\mathbf{E}}_n(\mathbf{r})$ 中介电常数张量 $\vec{\epsilon}_n^{(1)}$ 无关，只与等效标量介电常数 $\epsilon_n^{(1)}$ 与与之同向的相速度 v_n 是否匹配有关。其中 $\epsilon_n^{(1)}$ 来源于 $\mathbf{D}_n(\mathbf{r}) = \epsilon_n^{(1)} \mathbf{E}_{n\perp}(\mathbf{r})$ ，属于极化过程中的 $\epsilon_n^{(1)}$ ，由折射率椭球和极化方向决定；而 v_n 的大小由相速度曲面方程即折射率曲面和 v_n 的方向和决定，二者要想一一对应，唯一的解释是：

由于极化方向与波法线方向的夹角甚至相对位置，不是极化方向或波法线方向的函数，不随着光的传播方向的改变而改变（偏振 $\mathbf{E}_n(\mathbf{r})$ 方向与 \hat{k}_n 的夹角就可随着 \hat{k}_n 改变，对 e 光而言），并且决定 v_n 的大小的折射率与决定极化程度的折射率相同，均使用确定的波法线方向所同时确定下来的极化 $\mathbf{D}_n(\mathbf{r})$ 方向（ $\perp \hat{k}_n$ 的方向）上的折射率大小，即折射率曲面是表象，其本质上由折射率椭球决定，并且极化方向与波矢方向本质上是地位等同的同样有效的方向。

比如 o 光的折射率曲面是个球面，本质上是因为在 o 光波矢朝各个方向时，所对应的 o 光的极化/偏振方向上的折射率椭球所给出的等效标量折射率，随 o 光波矢的各个朝向而取值不变 $\epsilon_{no}^{(1)} = \epsilon_{n\perp}^{(1)}$ ，而不是因为 o 光的折射率椭球是个球面：o 光和 e 光同享一个旋转椭球面折射率椭球，只不过 o 光因其 $\mathbf{E}_{no}(\mathbf{r}), \mathbf{D}_{no}(\mathbf{r})$ 恒 $\perp \mathbf{k}$ 而恒极化于折射率椭球的上旋转轴方向，并取该方向的恒定折射率罢。这就将折射率椭球中决定极化大小且由极化方向决定的折射率，与折射率曲面中决定相速度大小，且由相速度方向决定的折射率，联系起来了。因此 o 光仍满足波动方程，波动方程允许 o 光的存在。

④. 一阶非线性系数各分量是实数时，e 光也是线偏振光，并可引入离散角概念
对于 e 光，还可考察其第二个方程，有

$$\frac{E_{nez}}{E_{ney}} = \frac{1 - \frac{\epsilon_{n\perp}^{(1)}}{\epsilon_{ne}^{(1)}}}{\left(1 - \frac{\epsilon_{n\parallel}^{(1)}}{\epsilon_{n\perp}^{(1)}}\right) k_{ne0y} k_{ne0z}} \quad (C121)$$

一方面据 e 光折射率曲面导出的

$$1 - \frac{\epsilon_{n\parallel}^{(1)}}{\epsilon_{n\perp}^{(1)}} = \frac{1}{k_{ne0z} k_{ne0z}} \left(1 - \frac{\epsilon_{n\parallel}^{(1)}}{\epsilon_{ne}^{(1)}}\right) \quad (C122)$$

代入得

$$\frac{E_{nez}}{E_{ney}} = \frac{1 - \frac{\epsilon_{n\perp}^{(1)}}{\epsilon_{ne}^{(1)}}}{\left(1 - \frac{\epsilon_{n//}^{(1)}}{\epsilon_{ne}^{(1)}}\right) \frac{k_{ne0y}}{k_{ne0z}}} = \frac{(\epsilon_{ne}^{(1)} - \epsilon_{n\perp}^{(1)}) k_{ne0z}}{(\epsilon_{ne}^{(1)} - \epsilon_{n//}^{(1)}) k_{ne0y}} \quad (C123)$$

据 e 光的折射率曲面 $\epsilon_{ne}^{(1)}(k_{ne0x}^2 + k_{ne0y}^2)/\epsilon_{n//}^{(1)} + \epsilon_{ne}^{(1)}k_{ne0z}^2/\epsilon_{n\perp}^{(1)} = 1$, 有

$$\begin{aligned} \left(\frac{\epsilon_{ne}^{(1)}}{\epsilon_{n\perp}^{(1)}} - 1\right) k_{ne0z}^2 &= \left(1 - \frac{\epsilon_{ne}^{(1)}}{\epsilon_{n//}^{(1)}}\right) (k_{ne0x}^2 + k_{ne0y}^2) \\ \frac{(\epsilon_{ne}^{(1)} - \epsilon_{n\perp}^{(1)}) k_{ne0z}}{(\epsilon_{ne}^{(1)} - \epsilon_{n//}^{(1)}) \sqrt{k_{ne0x}^2 + k_{ne0y}^2}} &= -\frac{\epsilon_{n\perp}^{(1)} \sqrt{k_{ne0x}^2 + k_{ne0y}^2}}{\epsilon_{n//}^{(1)} k_{ne0z}} \end{aligned} \quad (C124)$$

因 $k_{ne0x} = 0$, 代入得:

$$\begin{aligned} \cot(\theta'_{ne} + \frac{\pi}{2}) &:= \frac{E_{nez}}{E_{ney}} = \frac{(\epsilon_{ne}^{(1)} - \epsilon_{n\perp}^{(1)}) k_{ne0z}}{(\epsilon_{ne}^{(1)} - \epsilon_{n//}^{(1)}) k_{ne0y}} = -\frac{\epsilon_{n\perp}^{(1)} k_{ne0y}}{\epsilon_{n//}^{(1)} k_{ne0z}} = -\frac{\epsilon_{n\perp}^{(1)} \sin \theta_{ne}}{\epsilon_{n//}^{(1)} \cos \theta_{ne}} \\ &= -\frac{\epsilon_{n\perp}^{(1)}}{\epsilon_{n//}^{(1)}} \tan \theta_{ne} := -\tan \theta'_{ne} \end{aligned} \quad (C125)$$

从中可见, 当 $\epsilon_{n\perp}^{(1)}/\epsilon_{n//}^{(1)}$ 为一实数时, E_{nez}/E_{ney} 也是实的, 说明 e 光各分量间的比值是实数, 所以 e 光也是偏振光, 而不是一个分量相位提前, 一个分量相位滞后的圆或椭圆偏振光, 且 e 光的波动方程的解也不会含衰减因子。此时 θ'_{ne} 也为一实数, 因而具有意义。

另外, 利用式(C123), 式(C119)还可拓展为:

$$\widehat{\mathbf{k}}_{ne} \cdot \mathbf{E}_{ne}(\mathbf{r}) = \left(1 - \frac{\epsilon_{n//}^{(1)}}{\epsilon_{ne}^{(1)}}\right) \frac{E_{nez}}{k_{ne0z}} = \left(1 - \frac{\epsilon_{n\perp}^{(1)}}{\epsilon_{ne}^{(1)}}\right) \frac{E_{ney}}{k_{ne0y}} \quad (C126)$$

接着有:

$$\cot(\theta_{ne} + \frac{\pi}{2}) = \frac{D_{nez}}{D_{ney}} = \frac{\epsilon_{n//}^{(1)} E_{nez}}{\epsilon_{n\perp}^{(1)} E_{ney}} = -\frac{k_{ne0y}}{k_{ne0z}} = -\frac{\sin \theta_{ne}}{\cos \theta_{ne}} = -\tan \theta_{ne} \quad (C127)$$

于是得到离散角的正切值:

$$\tan \alpha_{ne} = \tan(\theta'_{ne} - \theta_{ne}) = \frac{\tan \theta'_{ne} - \tan \theta_{ne}}{1 + \tan \theta'_{ne} \cdot \tan \theta_{ne}} = \frac{\frac{\epsilon_{n\perp}^{(1)}}{\epsilon_{n//}^{(1)}} \tan \theta_{ne} - \tan \theta_{ne}}{1 + \frac{\epsilon_{n\perp}^{(1)}}{\epsilon_{n//}^{(1)}} \tan \theta_{ne} \cdot \tan \theta_{ne}} = \frac{\left(\frac{\epsilon_{n\perp}^{(1)}}{\epsilon_{n//}^{(1)}} - 1\right) \frac{1}{2} \sin 2\theta_{ne}}{\cos^2 \theta_{ne} + \frac{\epsilon_{n\perp}^{(1)}}{\epsilon_{n//}^{(1)}} \sin^2 \theta_{ne}}$$

$$= \frac{1}{2} \sin 2\theta_{ne} \left(\frac{1}{\varepsilon_{n\parallel}^{(1)}} - \frac{1}{\varepsilon_{n\perp}^{(1)}} \right) \cdot \frac{1}{\frac{\sin^2 \theta_{ne}}{\varepsilon_{n\parallel}^{(1)}} + \frac{\cos^2 \theta_{ne}}{\varepsilon_{n\perp}^{(1)}}} = \frac{1}{2} \sin 2\theta_{ne} \left(\frac{1}{\varepsilon_{n\parallel}^{(1)}} - \frac{1}{\varepsilon_{n\perp}^{(1)}} \right) \cdot \varepsilon_{ne}^{(1)} \quad (\text{C128})$$

⑤. $\tilde{\mathcal{H}}_n \perp \tilde{\mathcal{S}}_n, \mathbf{k}_n, \tilde{\mathcal{E}}_n, \tilde{\mathcal{D}}_n$ 的同时， $\tilde{\mathcal{E}}_n$ 偏离 $\tilde{\mathcal{D}}_n$ 的夹角，即 $\tilde{\mathcal{S}}_n$ 偏离 \mathbf{k}_n 的夹角

利用此结论，考虑同向于 $\mathcal{S}_n := \mathcal{S}_n = (\mathbf{E}_{0n} + \mathbf{E}_{0n}^*) \times (\mathbf{H}_{0n} + \mathbf{H}_{0n}^*) = 4\mathbf{E}_{0n} \times \mathbf{H}_{0n} =: 4\mathcal{S}_{0n}$ （注， \mathcal{S}_n 不含空，即没有 $\mathcal{S}_n(\mathbf{r})$ ）、 \mathcal{S}_{0n} 和 $\widehat{\mathcal{S}}_n = \mathcal{S}_{n0}$ 的 $\tilde{\mathcal{S}}_n = (\tilde{\mathbf{E}}_n + \tilde{\mathbf{E}}_n^*) \times (\tilde{\mathbf{H}}_n + \tilde{\mathbf{H}}_n^*) = \mathcal{S}_n \cos^2(\mathbf{k}_n \cdot \mathbf{r} - \omega_n t)$ 。由于已定义 $\tilde{\mathcal{E}}_n := \tilde{\mathbf{E}}_n + \tilde{\mathbf{E}}_n^*$ 、 $\tilde{\mathcal{H}}_n := \tilde{\mathbf{H}}_n + \tilde{\mathbf{H}}_n^*$ ，则 $\tilde{\mathcal{S}}_n = \tilde{\mathcal{E}}_n \times \tilde{\mathcal{H}}_n$ ，于是有 $\tilde{\mathcal{S}}_n \perp \tilde{\mathcal{H}}_n$ ；据(C94)的式二、式四，知 $\mathbf{k}_n, \mathcal{D}_n(\mathbf{r}) \perp \mathbf{H}_n(\mathbf{r})$ ，乘上时间项并加上复共轭项 ($\tilde{\mathbf{E}}_n^*$ 与 $\tilde{\mathbf{H}}_n^*$ 仍满足(C94)的式二、式四) 后因线性叠加仍有 $\mathbf{k}_n, \tilde{\mathcal{D}}_n \perp \tilde{\mathcal{H}}_n$ ；据(C94)的式三，知 $\mathbf{k}_n, \mathcal{E}_n(\mathbf{r}), \mathcal{D}_n(\mathbf{r})$ 三者共面，同理 $\mathbf{k}_n, \tilde{\mathcal{E}}_n, \tilde{\mathcal{D}}_n$ 三者共面，则 $\tilde{\mathcal{S}}_n, \mathbf{k}_n, \tilde{\mathcal{E}}_n, \tilde{\mathcal{D}}_n$ 四者共面，且 $\tilde{\mathcal{H}}_n \perp$ 此面。

再据(C94)的式四的含时实数形式 $\tilde{\mathcal{H}}_n = -v_n \tilde{\mathcal{D}}_n \times \widehat{\mathbf{k}}_n$ ，将 $\tilde{\mathcal{D}}_n \times$ 作用于两侧，得 $\tilde{\mathcal{D}}_n \times \tilde{\mathcal{H}}_n = v_n \tilde{\mathcal{D}}_n^2 \widehat{\mathbf{k}}_n$ ，它与 $\tilde{\mathcal{E}}_n \times \tilde{\mathcal{H}}_n = \tilde{\mathcal{S}}_n$ 相对应，可以说明 $\tilde{\mathcal{D}}_n \perp \mathbf{k}_n(\mathbf{r})$ 、 $\tilde{\mathcal{E}}_n \perp \tilde{\mathcal{S}}_n$ ，并且有：

$$\tilde{S}_{n\parallel} = \tilde{\mathcal{S}}_n \cdot \widehat{\mathbf{k}}_n = (\tilde{\mathcal{E}}_n \times \tilde{\mathcal{H}}_n) \cdot \widehat{\mathbf{k}}_n = (\tilde{\mathcal{H}}_n \times \widehat{\mathbf{k}}_n) \cdot \tilde{\mathcal{E}}_n \xrightarrow[(1-37 \text{ 左侧四式})]{\tilde{\mathcal{H}}_n \times \widehat{\mathbf{k}}_n = v_n \tilde{\mathcal{D}}_n} v_n \tilde{\mathcal{D}}_n \cdot \tilde{\mathcal{E}}_n = v_n \tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n \quad (\text{C129})$$

而对于 $\tilde{\mathcal{D}}_n \perp \mathbf{k}_n(\mathbf{r})$ 、 $\tilde{\mathcal{E}}_n \perp \tilde{\mathcal{S}}_n$ 还可进一步拓展有：

$$\begin{aligned} <\tilde{\mathcal{S}}_{ne}, \widehat{\mathbf{k}}_{ne}> &= <\tilde{\mathcal{S}}_{ne}, \mathbf{k}> - <\widehat{\mathbf{k}}_{ne}, \mathbf{k}> = \theta'_{ne} - \theta_{ne} \\ &= <\tilde{\mathcal{E}}_{ne}, \tilde{\mathcal{D}}_{ne}> = <\tilde{\mathcal{E}}_{ne}, \mathbf{k}> - <\tilde{\mathcal{D}}_{ne}, \mathbf{k}> = \left(\theta'_{ne} + \frac{\pi}{2} \right) - \left(\theta_{ne} + \frac{\pi}{2} \right) := \alpha_{ne} \end{aligned} \quad (\text{C130})$$

其中规定朝 \mathbf{k} 偏双正介质中的 $\widehat{\mathbf{k}}_{ne}$ 的方向选定 e 光坐标系的 \mathbf{j}_{ne} ，且假设双正介质中的 $\tilde{\mathcal{H}}_{ne} = v_{ne} \widehat{\mathbf{k}}_{ne} \times \tilde{\mathcal{D}}_{ne}$ 与 $\mathbf{i}_{ne} = \mathbf{j}_{ne} \times \mathbf{k}$ 反向。并将 $z-o-y$ 系的 y 轴作为二维直角坐标系的 y 轴， z 轴作为 x 轴，以搭配 cos 使用，注意这可能导致有点左手系的不适感，但方便描述。于是有 $0 << \widehat{\mathbf{k}}_{ne}, \mathbf{k}> = <\tilde{\mathcal{D}}_{ne}, \mathbf{j}_{ne}> := \theta_{ne} <\pi/2$ 、 $<\tilde{\mathcal{S}}_{ne}, \mathbf{k}> = <\tilde{\mathcal{E}}_{ne}, \mathbf{j}_{ne}> := \theta'_{ne}$ 。

⑥. 假设左手材料中的 $\tilde{\mathcal{E}}_{ne}, \widehat{\mathbf{k}}_{ne}, \mathbf{i}_{ne}, \mathbf{j}_{ne}, \mathbf{k}$ 与右手材料中的相同，则右左手正单轴晶体中的 $\tilde{\mathcal{S}}_{ne}$ 均比 \mathbf{k}_{ne} 更靠近光轴 \mathbf{k}

可见当 $\varepsilon_{ne}^{(1)} > 0$ 时，离散角 α_{ne} 同号于 $\varepsilon_{n\parallel}^{(1)} - \varepsilon_{n\perp}^{(1)}$ ，因此对于正单轴晶体 $\varepsilon_{n\parallel}^{(1)} > \varepsilon_{n\perp}^{(1)}$ ，而一般 $-\pi/2 < \alpha_{ne} < \pi/2$ ，则此时 $-\pi/2 < \alpha_{ne} < 0$ ，所以此时 e 光的 $\tilde{\mathcal{S}}_{ne}$ 比其 \mathbf{k}_{ne} 更靠近光轴；当 $\varepsilon_{ne}^{(1)} < 0$ 时，为方便考察，假设晶体中的光波的 $\tilde{\mathcal{E}}_{ne}, \widehat{\mathbf{k}}_{ne}, \mathbf{k}$ 与 $\varepsilon_{ne}^{(1)} > 0$ 时相同，则 $\tilde{\mathcal{D}}_{ne} = \varepsilon_{ne}^{(1)} \tilde{\mathcal{E}}_{ne\perp}$ ， $\tilde{\mathcal{H}}_{ne} = v_{ne} \widehat{\mathbf{k}}_{ne} \times \tilde{\mathcal{D}}_{ne}$ ， $\tilde{\mathcal{S}}_{ne} = \tilde{\mathcal{E}}_{ne} \times \tilde{\mathcal{H}}_{ne}$ 三者均反号， $\tilde{\mathcal{B}}_{ne}, \mathbf{i}_{ne}, \mathbf{j}_{ne} = \mathbf{k} \times \mathbf{i}_{ne}$ 不变，原(类)右手系 $\tilde{\mathcal{E}}_{ne}, \tilde{\mathcal{H}}_{ne}, \widehat{\mathbf{k}}_{ne}$ 、标准右手系 $\tilde{\mathcal{D}}_{ne}, \tilde{\mathcal{B}}_{ne}, \widehat{\mathbf{k}}_{ne}$ 分别变为(类)左手系、标准左手系。原(类)右手系 $\tilde{\mathcal{E}}_{ne}, \tilde{\mathcal{B}}_{ne}, \widehat{\mathbf{k}}_{ne}$ 、标准右手系 $\tilde{\mathcal{D}}_{ne}, \tilde{\mathcal{H}}_{ne}, \widehat{\mathbf{k}}_{ne}$ 、 $\tilde{\mathcal{E}}_{ne}, \tilde{\mathcal{H}}_{ne}, \tilde{\mathcal{S}}_{ne}$ 、 $\mathbf{i}_{ne}, \mathbf{j}_{ne}, \mathbf{k}$ 仍分别保持(类)

右手系、标准右手系。

因此时 $\tilde{\mathbf{S}}_{ne}$ 反向，有：

$$\begin{aligned}
 <\tilde{\mathbf{S}}_{ne}, \mathbf{k}> &= \text{原} <\tilde{\mathbf{S}}_{ne}, \mathbf{k}> + \pi = \text{原/现} <\tilde{\mathbf{E}}_{ne}, \mathbf{j}_{ne}> + \pi = \theta'_{ne}(|\mathcal{E}_{ne}^{(1)}|) + \pi := \theta'_{ne}(-|\mathcal{E}_{ne}^{(1)}|) \\
 <\tilde{\mathbf{D}}_{ne}, \mathbf{j}_{ne}> &= \text{原} <\tilde{\mathbf{D}}_{ne}, \mathbf{j}_{ne}> + \pi = \text{原/现} <\widehat{\mathbf{k}}_{ne}, \mathbf{k}> + \pi = \theta_{ne}(|\mathcal{E}_{ne}^{(1)}|) + \pi := \theta_{ne}(-|\mathcal{E}_{ne}^{(1)}|) + \pi \\
 <\tilde{\mathbf{S}}_{ne}, \widehat{\mathbf{k}}_{ne}> &= <\tilde{\mathbf{S}}_{ne}, \mathbf{k}> - <\widehat{\mathbf{k}}_{ne}, \mathbf{k}> = <\tilde{\mathbf{E}}_{ne}, \mathbf{j}_{ne}> + \pi - (<\tilde{\mathbf{D}}_{ne}, \mathbf{j}_{ne}> - \pi) \xrightarrow{-2\pi} <\tilde{\mathbf{E}}_{ne}, \tilde{\mathbf{D}}_{ne}> \\
 <\tilde{\mathbf{E}}_{ne}, \mathbf{k}> &= \text{原} <\tilde{\mathbf{E}}_{ne}, \mathbf{k}> = \theta'_{ne}(|\mathcal{E}_{ne}^{(1)}|) + \frac{\pi}{2} = \theta'_{ne}(-|\mathcal{E}_{ne}^{(1)}|) - \frac{\pi}{2} \\
 <\tilde{\mathbf{D}}_{ne}, \mathbf{k}> &= \text{原} <\tilde{\mathbf{D}}_{ne}, \mathbf{k}> + \pi = \theta_{ne}(|\mathcal{E}_{ne}^{(1)}|) + \frac{\pi}{2} + \pi = \theta_{ne}(-|\mathcal{E}_{ne}^{(1)}|) + \frac{3\pi}{2} \quad (C131)
 \end{aligned}$$

以及：

$$\begin{aligned}
 \frac{\pi}{2} <<\tilde{\mathbf{S}}_{ne}, \widehat{\mathbf{k}}_{ne}>> &= \alpha_{ne}(-|\mathcal{E}_{ne}^{(1)}|) = \theta'_{ne}(-|\mathcal{E}_{ne}^{(1)}|) - \theta_{ne}(-|\mathcal{E}_{ne}^{(1)}|) \\
 &= \theta'_{ne}(|\mathcal{E}_{ne}^{(1)}|) + \pi - \theta_{ne}(|\mathcal{E}_{ne}^{(1)}|) = \alpha_{ne}(|\mathcal{E}_{ne}^{(1)}|) + \pi <\frac{3\pi}{2} \quad (C132)
 \end{aligned}$$

现为考察 $\tilde{\mathbf{S}}_{ne}$ 与 \mathbf{k}_{ne} 谁更靠近光轴，不应用 $\alpha_{ne}(-|\mathcal{E}_{ne}^{(1)}|) = <\tilde{\mathbf{S}}_{ne}, \widehat{\mathbf{k}}_{ne}>$ 来衡量，因为此时 $\tilde{\mathbf{S}}_{ne}, \widehat{\mathbf{k}}_{ne}$ 的夹角已经超过了 $\pi/2$ ，则（仍）需用原

$$\begin{aligned}
 -\frac{\pi}{2} <<\tilde{\mathbf{S}}_{ne}, -\mathbf{k}>> - <\widehat{\mathbf{k}}_{ne}, \mathbf{k}> &= (<\tilde{\mathbf{S}}_{ne}, \mathbf{k}> - <-\mathbf{k}, \mathbf{k}>) - <\widehat{\mathbf{k}}_{ne}, \mathbf{k}> \\
 &= \theta'_{ne}(-|\mathcal{E}_{ne}^{(1)}|) - \pi - \theta_{ne}(-|\mathcal{E}_{ne}^{(1)}|) = \theta'_{ne}(|\mathcal{E}_{ne}^{(1)}|) - \theta_{ne}(|\mathcal{E}_{ne}^{(1)}|) = \alpha_{ne}(|\mathcal{E}_{ne}^{(1)}|) <\frac{\pi}{2} \quad (C133)
 \end{aligned}$$

来衡量。而由于此时在左手材料中的右手坐标系 $\mathbf{i}_{ne}, \mathbf{j}_{ne}, \mathbf{k}$ 下仍有

$$\frac{E_{nez}}{E_{ney}} = \cot <\tilde{\mathbf{E}}_{ne}, \mathbf{k}> = \cot[\theta'_{ne}(-|\mathcal{E}_{ne}^{(1)}|) - \frac{\pi}{2}] = -\tan[\theta'_{ne}(-|\mathcal{E}_{ne}^{(1)}|)] \quad (C134)$$

$$\frac{D_{nez}}{D_{ney}} = \cot <\tilde{\mathbf{D}}_{ne}, \mathbf{k}> = \cot[\theta_{ne}(-|\mathcal{E}_{ne}^{(1)}|) + \frac{3\pi}{2}] = -\tan[\theta_{ne}(-|\mathcal{E}_{ne}^{(1)}|)] \quad (C134)$$

成立，于是

$$\tan \alpha_{ne}(-|\mathcal{E}_{ne}^{(1)}|) = \frac{1}{2} \sin[2\theta_{ne}(-|\mathcal{E}_{ne}^{(1)}|)] \cdot \left(\frac{1}{\mathcal{E}_{n//}^{(1)}} - \frac{1}{\mathcal{E}_{n\perp}^{(1)}} \right) \cdot \mathcal{E}_{ne}^{(1)} \quad (C134)$$

也成立，则

$$\begin{aligned}
 \tan \alpha_{ne}(|\mathcal{E}_{ne}^{(1)}|) &= \tan[\alpha_{ne}(-|\mathcal{E}_{ne}^{(1)}|) - \pi] = \tan \alpha_{ne}(-|\mathcal{E}_{ne}^{(1)}|) \\
 &= \frac{1}{2} \sin[2\theta_{ne}(|\mathcal{E}_{ne}^{(1)}|)] \cdot \left(\frac{1}{\mathcal{E}_{n//}^{(1)}} - \frac{1}{\mathcal{E}_{n\perp}^{(1)}} \right) \cdot \mathcal{E}_{ne}^{(1)} \quad (C135)
 \end{aligned}$$

因此仍有 $\alpha_{ne}(|\varepsilon_{ne}^{(1)}|)$ 同号于 $(\varepsilon_{n\perp}^{(1)} - \varepsilon_{n//}^{(1)}) \cdot \varepsilon_{ne}^{(1)}$ 。但注意虽然 $\alpha_{ne}(-|\varepsilon_{ne}^{(1)}|)$ 、 $\alpha_{ne}(|\varepsilon_{ne}^{(1)}|)$ 二者的正切值同号甚至相同，二者同号，二者的值也不相同： $\alpha_{ne}(-|\varepsilon_{ne}^{(1)}|) = \alpha_{ne}(|\varepsilon_{ne}^{(1)}|) + \pi$ 。

定义 $|\varepsilon_{n//}^{(1)}| > |\varepsilon_{n\perp}^{(1)}|$ 的折射率为负 $\varepsilon_{ne}^{(1)} < 0$ 的人工晶体为正单轴（假设 $\varepsilon_n^{(1)}$ 的各分量如 $\varepsilon_{n//}^{(1)}, \varepsilon_{n\perp}^{(1)}$ 都是负的，同时各个方向上的等效标量 $\varepsilon_n^{(1)}$ 也是负的），则此时 $(\varepsilon_{n\perp}^{(1)} - \varepsilon_{n//}^{(1)})$ 为正， $\varepsilon_{ne}^{(1)}$ 为负，有 $-\pi/2 < \alpha_{ne}(|\varepsilon_{ne}^{(1)}|) < 0$ 。此时 e 光的 \tilde{S}_{ne} 仍比其 k_{ne} 更靠近光轴。

因此无论是人工双负材料还是自然材料，由于 α_{ne} （包含 $\varepsilon_{ne}^{(1)} > 0$ 时的 $\alpha_{ne}(|\varepsilon_{ne}^{(1)}|)$ 和 $\varepsilon_{ne}^{(1)} < 0$ 时的 $\alpha_{ne}(-|\varepsilon_{ne}^{(1)}|)$ ）恒同号于 $(\varepsilon_{n\perp}^{(1)} - \varepsilon_{n//}^{(1)}) \cdot \varepsilon_{ne}^{(1)}$ ，所以在正单轴晶体中 e 光的 \tilde{S}_{ne} 均比其 k_{ne} 更靠近光轴，无论是双负还是双正材料。

⑦. 最大离散角及其应用：正入射时实现 oe 光能流分离程度最大的能流双折射现考察

$$\begin{aligned} \frac{d\alpha_{ne}}{d\theta_{ne}} &= \frac{d\theta'_{ne}}{d\theta_{ne}} - 1 = \frac{d}{d\theta_{ne}} \arctan \left(\frac{\varepsilon_{n\perp}^{(1)}}{\varepsilon_{n//}^{(1)}} \tan \theta_{ne} \right) - 1 = \frac{1}{1 + \left(\frac{\varepsilon_{n\perp}^{(1)}}{\varepsilon_{n//}^{(1)}} \tan \theta_{ne} \right)^2} \frac{\varepsilon_{n\perp}^{(1)}}{\varepsilon_{n//}^{(1)}} \sec^2 \theta_{ne} - 1 \\ &= \frac{1}{\frac{\sin^2 \theta_{ne}}{\left[\varepsilon_{n//}^{(1)} \right]^2} + \frac{\cos^2 \theta_{ne}}{\left[\varepsilon_{n\perp}^{(1)} \right]^2}} \frac{1}{\varepsilon_{n//}^{(1)} \varepsilon_{n\perp}^{(1)}} - 1 = 0 \end{aligned} \quad (\text{C136})$$

得极值点

$$\begin{aligned} \frac{\sin^2 \theta_{ne}}{\left[\varepsilon_{n//}^{(1)} \right]^2} + \frac{\cos^2 \theta_{ne}}{\left[\varepsilon_{n\perp}^{(1)} \right]^2} &= \frac{1}{\varepsilon_{n//}^{(1)} \varepsilon_{n\perp}^{(1)}} \\ \left[\varepsilon_{n\perp}^{(1)} \right]^2 \sin^2 \theta_{ne} + \left[\varepsilon_{n//}^{(1)} \right]^2 \cos^2 \theta_{ne} &= \varepsilon_{n//}^{(1)} \varepsilon_{n\perp}^{(1)} \left[\sin^2 \theta_{ne} + \cos^2 \theta_{ne} \right] \\ \varepsilon_{n\perp}^{(1)} \left[\varepsilon_{n\perp}^{(1)} - \varepsilon_{n//}^{(1)} \right] \sin^2 \theta_{ne} &= \varepsilon_{n//}^{(1)} \left[\varepsilon_{n\perp}^{(1)} - \varepsilon_{n//}^{(1)} \right] \cos^2 \theta_{ne} \\ \tan^2 \theta_{ne} &= \frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}} \end{aligned} \quad (\text{C137})$$

代入得 α_{ne} 的极大值或极小值

$$\begin{aligned} \alpha_{neM} &= \arctan \left(\frac{\frac{\varepsilon_{n\perp}^{(1)}}{\varepsilon_{n//}^{(1)}} \tan \theta_{ne} - \tan \theta_{ne}}{1 + \frac{\varepsilon_{n\perp}^{(1)}}{\varepsilon_{n//}^{(1)}} \tan \theta_{ne} \cdot \tan \theta_{ne}} \right) = \arctan \left(\frac{\varepsilon_{ne}^{(1)}}{\left| \varepsilon_{ne}^{(1)} \right|} \frac{1}{2} \left[\sqrt{\frac{\varepsilon_{n\perp}^{(1)}}{\varepsilon_{n//}^{(1)}}} - \sqrt{\frac{\varepsilon_{n//}^{(1)}}{\varepsilon_{n\perp}^{(1)}}} \right] \right) = \\ \frac{\varepsilon_{ne}^{(1)}}{\left| \varepsilon_{ne}^{(1)} \right|} \arctan \left(\frac{\varepsilon_{n\perp}^{(1)} - \varepsilon_{n//}^{(1)}}{2 \sqrt{\varepsilon_{n//}^{(1)} \varepsilon_{n\perp}^{(1)}}} \right) &= \frac{\varepsilon_{ne}^{(1)}}{\left| \varepsilon_{ne}^{(1)} \right|} \arctan \left(\frac{n_{ne\perp}^2 - n_{ne//}^2}{2 n_{ne//} n_{ne\perp}} \right) \end{aligned} \quad (\text{C138})$$

其中为使 α_{neM} 与 α_{ne} 均始终同号于 $(\varepsilon_{n\perp}^{(1)} - \varepsilon_{n//}^{(1)}) \cdot \varepsilon_{ne}^{(1)}$ ，极值点需为

$$\tan \theta_{ne} = \pm \frac{n_{ne//}}{n_{ne\perp}} = \operatorname{sgn}(\varepsilon_{ne}^{(1)}) \frac{n_{ne//}}{n_{ne\perp}} = \frac{\varepsilon_{ne}^{(1)}}{|\varepsilon_{ne}^{(1)}|} \frac{n_{ne//}}{n_{ne\perp}} \quad (\text{C139})$$

在实际应用中，一般调整好晶面与光轴的夹角 β （如切削），使得 $\pi/2 - \beta_{ne} = \alpha_{neM}$ 后，将入射光垂直于晶面正入射晶体，此时根据折射定律，o 光 e 光的波法线方向相同且垂直于晶面；又 o 光的光线/波射线方向 $\widehat{\mathbf{S}_{no}}$ 同向于 o 光波法线方向，则此时 e 光的光线方向 $\widehat{\mathbf{S}_{ne}}$ 因与 $\widehat{\mathbf{k}_{ne}}$ 夹角最大，而与 $\widehat{\mathbf{S}_{no}}$ 夹角最大；可见双折射本身不仅针对波阵面或波法线，也针对能流或波线，且即使没有发生相速度的双折射，也可能发生能流的双折射。并且之后会给出，所有的折射、双折射均是指能流的折射、双折射，而不是波矢的折射、双折射。因此人们常说 e 光的折射不符合 Snell 定律或折射定律，实际上言下之意是 e 光的能流方向不能由 Snell 定律确定，但 e 光的波矢是符合 Snell 定律的。

附C1.2.9.6 菲涅尔波法线方程及其导出的两个光线曲面和三者朝双负材料中的拓展

(1) 菲涅尔波法线方程及其于双负材料中的拓展形式（波法线方程可类似拓展）

因 $\mathbf{S}_n, \mathbf{k}_n, \mathbf{E}_n(\mathbf{r}), \mathbf{D}_n(\mathbf{r})$ 四者共面，(C94)式三可写为

$$\begin{aligned} D_n \operatorname{sgn}(\varepsilon_n^{(1)}) \widehat{\mathbf{E}_{n\perp}} &= D_n \widehat{\mathbf{D}_n} = \mathbf{D}_n(\mathbf{r}) = \varepsilon_n^{(1)} \mathbf{E}_{n\perp}(\mathbf{r}) \\ &= \varepsilon_n^{(1)} E_n \cos[\alpha_n(|\varepsilon_n^{(1)}|)] \widehat{\mathbf{E}_{n\perp}} = \varepsilon_n^{(1)} E_n \cos \alpha_n \widehat{\mathbf{D}_n} \end{aligned} \quad (\text{C140})$$

于是

$$E_n = \frac{D_n}{\varepsilon_n^{(1)} \cos \alpha_n} = \frac{D_n \operatorname{sgn}(\varepsilon_n^{(1)})}{\varepsilon_n^{(1)} \cos [\alpha_n(|\varepsilon_n^{(1)}|)]} \quad (\text{C141})$$

代入有：

$$\begin{aligned} \mathbf{E}_n(\mathbf{r}) &= E_n \widehat{\mathbf{E}_n} = E_n \operatorname{sgn}(\varepsilon_n^{(1)}) \widehat{\mathbf{D}_{n\perp}} = \frac{D_n}{\varepsilon_n^{(1)} \cos [\alpha_n(|\varepsilon_n^{(1)}|)]} \widehat{\mathbf{D}_{n\perp}} \\ &= \frac{D_n \cos [\alpha_n(|\varepsilon_n^{(1)}|)]}{\varepsilon_n^{(1)} \cos^2 [\alpha_n(|\varepsilon_n^{(1)}|)]} \widehat{\mathbf{D}_{n\perp}} = \frac{\operatorname{sgn}(\varepsilon_{nr}^{(1)})}{\varepsilon_0 (n_{ne} \cos [\alpha_n(|\varepsilon_n^{(1)}|)])^2} \cdot \mathbf{D}_{n\perp}(\mathbf{r}) \\ &:= \frac{1}{\varepsilon_0 n_{ner} |n_{ner}|} [\mathbf{D}_n(\mathbf{r}) - \mathbf{D}_{n//}(\mathbf{r})] = \frac{1}{\varepsilon_0 n_{ner} |n_{ner}|} \{ \mathbf{D}_n(\mathbf{r}) - [\widehat{\mathbf{S}_n} \cdot \mathbf{D}_n(\mathbf{r})] \widehat{\mathbf{S}_n} \} \end{aligned} \quad (\text{C142})$$

其中可有 $\varepsilon_n^{(1)} < 0$ ，对应 $\pi/2 < \alpha_n(-|\varepsilon_n^{(1)}|) < 3\pi/2$ 、 $-\pi/2 < \alpha_n(|\varepsilon_n^{(1)}|) < \pi/2$ ，此时 $n_{ne}, n_{ner} < 0$ ；

其中 $\mathbf{D}_n(\mathbf{r}), \mathbf{E}_{n\perp}(\mathbf{r}) \perp \widehat{\mathbf{k}}_n$ 、 $\mathbf{E}_n(\mathbf{r}), \mathbf{D}_{n\perp}(\mathbf{r}) \perp \widehat{\mathbf{S}}_n$ ，但每对内部二者均可反向。而类比

$$v_{np} := v_n = \frac{c}{|n_n|} = \frac{c}{|n_{nm}| |n_{ne}|} \quad (C143)$$

有：

$$\begin{aligned} v_{nr} &:= \frac{c}{|n_{nm}| |n_{ner}|} = \frac{c}{|n_{nm}| |n_{ne}| \cos[\alpha_n(|\epsilon_n^{(1)}|)]} = \frac{v_{np}}{\cos[\alpha_n(|\epsilon_n^{(1)}|)]} > 0 \\ \mathbf{v}_{nr} &:= v_{nr} \widehat{\mathbf{S}}_n = v_{nr} \widehat{\mathbf{v}}_{nr} \end{aligned} \quad (C144)$$

将式(C142)写成分量形式即

$$\begin{aligned} E_{nx} &= \frac{1}{\epsilon_0 n_{ner} |n_{ner}|} \left\{ D_{nx} - [\widehat{\mathbf{S}}_n \cdot \mathbf{D}_n(\mathbf{r})] S_{n0x} \right\} = \frac{1}{\epsilon_0 n_{ner} |n_{ner}|} \cdot \left\{ \epsilon_{nx}^{(1)} E_{nx} - [\widehat{\mathbf{S}}_n \cdot \mathbf{D}_n(\mathbf{r})] S_{n0x} \right\} \\ &= \frac{1}{n_{ner} |n_{ner}|} \left\{ \epsilon_{rnx}^{(1)} E_{nx} - \frac{1}{\epsilon_0} [\widehat{\mathbf{S}}_n \cdot \mathbf{D}_n(\mathbf{r})] S_{n0x} \right\} \end{aligned} \quad (C145)$$

得到

$$\begin{aligned} E_{nx} &= -\frac{1}{\epsilon_0} [\widehat{\mathbf{S}}_n \cdot \mathbf{D}_n(\mathbf{r})] \cdot \frac{S_{n0x}}{n_{ner} |n_{ner}| - n_{nex} |n_{nex}|} = -\frac{|\mu_{rn}^{(1)}|}{\epsilon_0} [\widehat{\mathbf{S}}_n \cdot \mathbf{D}_n(\mathbf{r})] \frac{S_{n0x}}{n_{nr} |n_{nr}| - n_{nx} |n_{nx}|} \\ &\xrightarrow{\epsilon_{ner}^{(1)}, \epsilon_{nex}^{(1)} > 0} -\frac{\mu_{rn}^{(1)}}{\epsilon_0} [\widehat{\mathbf{S}}_n \cdot \mathbf{D}_n(\mathbf{r})] \frac{S_{n0x}}{n_{nr}^2 - n_{nx}^2} \end{aligned} \quad (C146)$$

同理其他三分量。于是

$$\begin{aligned} \mathbf{E}_n(\mathbf{r}) \cdot \widehat{\mathbf{S}}_n &= E_{nx} S_{n0x} + E_{ny} S_{n0y} + E_{nz} S_{n0z} = \\ -\frac{1}{\epsilon_0} [\widehat{\mathbf{S}}_n \cdot \mathbf{D}_n(\mathbf{r})] &\left[\frac{S_{n0x}^2}{\epsilon_{rnr}^{(1)} - \epsilon_{rnx}^{(1)}} + \frac{S_{n0y}^2}{\epsilon_{rnr}^{(1)} - \epsilon_{rny}^{(1)}} + \frac{S_{n0z}^2}{\epsilon_{rnr}^{(1)} - \epsilon_{rnz}^{(1)}} \right] = 0 \end{aligned} \quad (C147)$$

而一般 $\widehat{\mathbf{S}}_n \cdot \mathbf{D}_n(\mathbf{r}) \neq 0$ ，得到

$$\frac{S_{n0x}^2}{\epsilon_{rnr}^{(1)} - \epsilon_{rnx}^{(1)}} + \frac{S_{n0y}^2}{\epsilon_{rnr}^{(1)} - \epsilon_{rny}^{(1)}} + \frac{S_{n0z}^2}{\epsilon_{rnr}^{(1)} - \epsilon_{rnz}^{(1)}} = 0 \quad (C148)$$

或写作

$$\frac{S_{n0x}^2}{n_{ner} |n_{ner}| - n_{nex} |n_{nex}|} + \frac{S_{n0y}^2}{n_{ner} |n_{ner}| - n_{ney} |n_{ney}|} + \frac{S_{n0z}^2}{n_{ner} |n_{ner}| - n_{nez} |n_{nez}|} = 0 \quad (C149)$$

当 $\epsilon_{rnr}^{(1)}$ 与各分量 $\epsilon_{rnx}^{(1)}, \epsilon_{rny}^{(1)}, \epsilon_{rnz}^{(1)}$ 均同号时，即有菲涅尔光线/波射线方程：

$$\frac{S_{n0x}^2}{n_{ner}^2 - n_{nex}^2} + \frac{S_{n0y}^2}{n_{ner}^2 - n_{ney}^2} + \frac{S_{n0z}^2}{n_{ner}^2 - n_{nez}^2} = 0 \quad (C150)$$

同理菲涅尔波法线方程也有类似的拓展，其第二种得来方法就可视为将其拓展到对于 $\varepsilon_n^{(1)}$ 或某些分量为负也同样适用，但要注意相应的 D_{ni} 等物理量也需被同步拓展。

(2) 通过菲涅尔波射线方程得出 o 光 e 光光波面及其在双负材料中的拓展形式对比式(C102)所得到的式(C104)，对光线方程有类似的

$$\begin{aligned} & \left[\left(\frac{S_{n0x}}{n_{ner}} \right)^2 + \left(\frac{S_{n0y}}{n_{ner}} \right)^2 + \left(\frac{S_{n0z}}{n_{ner}} \right)^2 - \frac{1}{n_{nex}^2} \right] \cdot \left[\frac{1}{n_{nex}^2} \left(\frac{S_{n0x}}{n_{ner}} \right)^2 + \frac{1}{n_{ney}^2} \left(\frac{S_{n0y}}{n_{ner}} \right)^2 + \frac{1}{n_{nez}^2} \left(\frac{S_{n0z}}{n_{ner}} \right)^2 - \frac{1}{n_{ney}^2} \frac{1}{n_{nez}^2} \right] \\ & + \left[\frac{1}{n_{nex}^2} - \frac{1}{n_{ney}^2} \right] \cdot \left[\frac{1}{n_{nex}^2} - \frac{1}{n_{nez}^2} \right] \cdot \left(\frac{S_{n0x}}{n_{ner}} \right)^2 = 0 \end{aligned} \quad (C151)$$

将 $n_{nex} = n_{ney} =: n_{ne\perp}$ 、 $n_{nez} =: n_{ne//}$ 代入其中，得到

$$\begin{aligned} & \left[\left(\frac{S_{n0x}}{n_{ner}} \right)^2 + \left(\frac{S_{n0y}}{n_{ner}} \right)^2 + \left(\frac{S_{n0z}}{n_{ner}} \right)^2 - \frac{1}{n_{ne\perp}^2} \right] \\ & \cdot \left[\frac{1}{n_{ne\perp}^2} \left(\frac{S_{n0x}}{n_{ner}} \right)^2 + \frac{1}{n_{ne\perp}^2} \left(\frac{S_{n0y}}{n_{ner}} \right)^2 + \frac{1}{n_{ne\perp}^2} \left(\frac{S_{n0z}}{n_{ner}} \right)^2 - \frac{1}{n_{ne\perp}^2} \frac{1}{n_{ne//}^2} \right] = 0 \end{aligned} \quad (C152)$$

写成三角函数形式即

$$n_{nero}^2 = n_{ne\perp}^2 \quad n_{nere}^2 = n_{ne//}^2 \sin^2 \theta'_{ne} + n_{ne\perp}^2 \cos^2 \theta'_{ne} \quad (C153)$$

对于 $\varepsilon_n^{(1)} > 0$ ，得 o 光 e 光的光波面/光线曲面：

$$n_{nero} = n_{ne\perp} \quad n_{nere}(\theta'_{ne}) = \sqrt{n_{ne//}^2 \sin^2 \theta'_{ne} + n_{ne\perp}^2 \cos^2 \theta'_{ne}} \quad (C154)$$

更一般地，有

$$\varepsilon_{rno}^{(1)} = \varepsilon_{rn\perp}^{(1)} \quad \varepsilon_{rne}^{(1)} = \varepsilon_{rn//}^{(1)} \left[S_{ne0x}^2 + S_{ne0y}^2 \right] + \varepsilon_{rn\perp}^{(1)} S_{ne0z}^2 = \varepsilon_{rn//}^{(1)} \sin^2 \theta'_{ne} + \varepsilon_{rn\perp}^{(1)} \cos^2 \theta'_{ne} \quad (C155)$$

两侧也可同乘以 ε_0 以变得更简洁。

进一步地，当 $\varepsilon_{no}^{(1)} < 0$ 时，此时 $n_{nex} = n_{ney} =: n_{ne\perp} < 0$ ，仍有 $n_{nero} = n_{ne\perp}$ ；当 $\varepsilon_{ne}^{(1)} < 0$ 时，
 $n_{nere}^2 = -\varepsilon_{rne}^{(1)} = -\varepsilon_{rn//}^{(1)} \sin^2 \theta'_{ne} - \varepsilon_{rn\perp}^{(1)} \cos^2 \theta'_{ne}$ ，
 $n_{nere}(\theta'_{ne}) = -\sqrt{-\varepsilon_{rn//}^{(1)} \sin^2 \theta'_{ne} - \varepsilon_{rn\perp}^{(1)} \cos^2 \theta'_{ne}}$ ，当
 $\varepsilon_{n//}^{(1)}, \varepsilon_{n\perp}^{(1)} < 0$ 时， $n_{nere}(\theta'_{ne}) = -\sqrt{n_{ne//}^2 \sin^2 \theta'_{ne} + n_{ne\perp}^2 \cos^2 \theta'_{ne}}$ 。

附C1.2.9.7 双正、双负线性材料中场能密度、能流密度的表达式

(1) 解为平面波的情形下，线性介质中，电场能密度恒等于磁场能密度，即使电磁均各向异性

下面考察能流密度本身的性质。在线性介质中，据式(附 D11-13)，有

$$\tilde{w} = \tilde{w}_e + \tilde{w}_m = \frac{1}{2} \tilde{\mathcal{E}} \cdot \tilde{\mathcal{D}} + \frac{1}{2} \tilde{\mathcal{H}} \cdot \tilde{\mathcal{B}} \quad (\text{C156})$$

两边各量单色化得

$$\tilde{w}_n = \frac{1}{2} \tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n + \frac{1}{2} \tilde{\mathcal{H}}_n \cdot \tilde{\mathcal{B}}_n \quad (\text{C157})$$

代入(C93)的左侧四式 $\tilde{\mathcal{H}}_n \times \hat{\mathbf{k}}_n = v_n \tilde{\mathcal{D}}_n$ 和三式 $\tilde{\mathcal{E}}_n \times \hat{\mathbf{k}}_n = -v_n \tilde{\mathcal{B}}_n$ ，得

$$\tilde{w}_n = \frac{1}{2v_n} \tilde{\mathcal{E}}_n \cdot [\tilde{\mathcal{H}}_n \times \hat{\mathbf{k}}_n] + \frac{1}{2v_n} \tilde{\mathcal{H}}_n \cdot [\hat{\mathbf{k}}_n \times \tilde{\mathcal{E}}_n] = \frac{1}{v_n} \hat{\mathbf{k}}_n \cdot [\tilde{\mathcal{E}}_n \times \tilde{\mathcal{H}}_n] = \frac{1}{v_n} \hat{\mathbf{k}}_n \cdot \tilde{\mathcal{S}}_n = \frac{1}{v_n} \tilde{S}_{n\parallel} \quad (\text{C158})$$

同时也说明即使在电磁均各向异性的线性介质中，只要解为平面电磁波，则(C93)四式可取等，则上述关系成立，并有：

$$\tilde{w}_e = \tilde{w}_m \quad , \quad \tilde{\mathcal{E}} \cdot \tilde{\mathcal{D}} = \tilde{\mathcal{H}} \cdot \tilde{\mathcal{B}} \quad (\text{C159})$$

而在磁各向同性的线性介质中，由于解一定为定振幅平面波，则上述关系恒成立。

(2) 能流密度及其平行分量和垂直分量表达式

将式(C159)与之前的结论 $\tilde{S}_{n\parallel} = \tilde{\mathcal{S}}_n \cdot \hat{\mathbf{k}}_n = v_n \tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n$ 对比，或将 $\tilde{w}_{en} = \tilde{w}_{mn}$ 代入 \tilde{w}_n 得：

$$\tilde{S}_{n\parallel} = \tilde{\mathcal{S}}_n \cdot \hat{\mathbf{k}}_n = \tilde{\mathcal{S}}_n \cos \alpha_n = \tilde{w}_n v_n = 2 \tilde{w}_{en} v_n = 2 \tilde{w}_{mn} v_n \quad (\text{C160})$$

$$\tilde{S}_n = \frac{\tilde{S}_{n\parallel}}{\cos \alpha_n} = \tilde{w}_n \frac{v_n}{\cos \alpha_n} = \tilde{w}_n \frac{v_n}{\operatorname{sgn}(\epsilon_n^{(1)}) \cdot \cos [\alpha_n(|\epsilon_n^{(1)}|)]} = |\tilde{w}_n| v_{nr} \quad (\text{C161})$$

$$\tilde{\mathcal{S}}_n = |\tilde{w}_n| v_{nr} \quad (\text{C162})$$

$\tilde{\mathcal{S}}_n$ 还有以下性质：

$$|\tilde{S}_{n\perp}| = |\hat{\mathbf{k}}_n \times \tilde{\mathcal{S}}_n| = |\hat{\mathbf{k}}_n \times (\tilde{\mathcal{E}}_n \times \tilde{\mathcal{H}}_n)| = |(\tilde{\mathcal{H}}_n \cdot \hat{\mathbf{k}}_n) \tilde{\mathcal{E}}_n - (\tilde{\mathcal{E}}_n \cdot \hat{\mathbf{k}}_n) \tilde{\mathcal{H}}_n| = |-(\tilde{\mathcal{E}}_n \cdot \hat{\mathbf{k}}_n) \tilde{\mathcal{H}}_n| = |\tilde{\mathcal{E}}_{n\parallel} \tilde{\mathcal{H}}_n|$$

$$\tilde{S}_{n\perp} = -\tilde{\mathcal{E}}_{n\parallel} \tilde{\mathcal{H}}_n \quad ^\textcircled{1} \quad (\text{C163})$$

^① 由于 $\mathbf{k}_n, \mathcal{D}_n$ 相互正交，因此 $\mathcal{S}_n, \mathcal{E}_n$ 均参照该坐标系，即有 $S_{n\parallel} := \mathcal{S}_n \cdot \hat{\mathbf{k}}_n$ 、 $S_{n\perp} := \mathcal{S}_n \cdot \widehat{\mathcal{D}}_n$ ，以及配套的 $\mathcal{S}_{n\parallel} = S_{n\parallel} \hat{\mathbf{k}}_n$ 、 $\mathcal{S}_{n\perp} = S_{n\perp} \widehat{\mathcal{D}}_n$ ，同理 $\mathcal{E}_{n\parallel}, \mathcal{E}_{n\perp}; \mathcal{E}_{n\parallel}, \mathcal{E}_{n\perp}$ ；同理因 $\mathcal{S}_n, \mathcal{E}_n$ 相互正交，因此 $\mathbf{k}_n, \mathcal{D}_n$ 参照此坐标系，即有 $\mathcal{D}_{n\parallel} := \mathcal{D}_n \cdot \widehat{\mathcal{S}}_n$ 、 $\mathcal{D}_{n\perp} := \mathcal{D}_n \cdot \widehat{\mathcal{E}}_n$ 以及 $\mathcal{D}_{n\parallel} = \mathcal{D}_{n\parallel} \widehat{\mathcal{S}}_n$ 、 $\mathcal{D}_{n\perp} = \mathcal{D}_{n\perp} \widehat{\mathcal{E}}_n$ ，同理 $k_{n\parallel}, k_{n\perp}; \mathbf{k}_{n\parallel}, \mathbf{k}_{n\perp}$ 。因此对于这四者一般不采用自己的基矢作为基矢，而采取对方的，如一般没有 $\mathcal{S}_{n\parallel} = S_{n\parallel} \widehat{\mathcal{S}}_{n\parallel}$ 、 $\mathcal{S}_{n\perp} = S_{n\perp} \widehat{\mathcal{S}}_{n\perp}$ ；另外，也因为没有采用自己的基矢作为基矢，而采取另一套正交基，这个原因，导致 $S_{n\parallel}, S_{n\perp}, \mathcal{E}_{n\parallel}, \mathcal{E}_{n\perp}; k_{n\parallel}, k_{n\perp}, \mathcal{D}_{n\parallel}, \mathcal{D}_{n\perp}$ 均可为负，以及 $\widehat{\mathcal{S}}_{n\parallel} = \hat{\mathbf{k}}_n$ 、 $\widehat{\mathcal{S}}_{n\perp} = \widehat{\mathcal{D}}_n$ 、 $\widehat{\mathcal{D}}_{n\parallel} = \widehat{\mathcal{S}}_n$ 、 $\widehat{\mathcal{D}}_{n\perp} = \widehat{\mathcal{E}}_n$ 等。

由于在双负或双正的正单轴晶体中，均有 e 光的 $\tilde{\mathbf{S}}_n$ 比 $\hat{\mathbf{k}}_n$ 更偏向光轴 \mathbf{k} ，即 $\alpha_n(\varepsilon_n^{(1)}) < 0$ ，则同样 $\tilde{\mathcal{E}}_n$ 比 $\tilde{\mathcal{D}}_n$ 更偏向 \mathbf{j}_n ，于是在这样的晶体中恒有 $\tilde{\mathcal{E}}_{n\parallel} > 0$ 、 $\tilde{\mathcal{S}}_{n\perp} < 0$ 。其中据式(C115)有

$$\tilde{\mathcal{E}}_{n\parallel} = \left(1 - \frac{\varepsilon_{n\parallel}^{(1)}}{\varepsilon_{n\perp}^{(1)}}\right) k_{n0z} \tilde{\mathcal{E}}_{nz} \quad (C164)$$

对 e 光，据式(C119)、式(C126)还有

$$\tilde{\mathcal{E}}_{ne\parallel} = \left(1 - \frac{\varepsilon_{ne\parallel}^{(1)}}{\varepsilon_{ne}^{(1)}}\right) \frac{\tilde{\mathcal{E}}_{nez}}{k_{ne0z}} = \left(1 - \frac{\varepsilon_{n\perp}^{(1)}}{\varepsilon_{ne}^{(1)}}\right) \frac{\tilde{\mathcal{E}}_{ney}}{k_{ne0y}} \quad (C165)$$

且在磁各向同性的线性介质中有

$$\begin{aligned} \tilde{\mathbf{S}}_n &= \tilde{\mathcal{E}}_n \times \tilde{\mathcal{H}}_n - \frac{\tilde{\mathcal{E}}_n \times \hat{\mathbf{k}}_n = -v_n \tilde{\mathbf{E}}_n}{1-37 \text{ 左侧三式}} = \frac{1}{v_n \mu_n^{(1)}} \tilde{\mathcal{E}}_n \times [\hat{\mathbf{k}}_n \times \tilde{\mathcal{E}}_n] = v_n \varepsilon_n^{(1)} [\tilde{\mathcal{E}}_n^2 \hat{\mathbf{k}}_n - \tilde{\mathcal{E}}_{n\parallel} \tilde{\mathcal{E}}_n] \\ &= \varepsilon_n^{(1)} (\tilde{\mathcal{E}}_{n\perp}^2 + \tilde{\mathcal{E}}_{n\parallel}^2) \mathbf{v}_n - v_n \varepsilon_n^{(1)} \tilde{\mathcal{E}}_{n\parallel} \tilde{\mathcal{E}}_n = \tilde{\mathcal{E}}_{n\perp} \tilde{\mathcal{D}}_n \mathbf{v}_n + \varepsilon_n^{(1)} \tilde{\mathcal{E}}_{n\parallel} v_n \left[\tilde{\mathcal{E}}_{n\parallel} \hat{\mathbf{k}}_n - \tilde{\mathcal{E}}_n \right] = \tilde{\mathcal{E}}_{n\perp} \tilde{\mathcal{D}}_n \mathbf{v}_n - \varepsilon_n^{(1)} \tilde{\mathcal{E}}_{n\parallel} v_n \tilde{\mathcal{E}}_{n\perp} \\ &= \tilde{\mathcal{E}}_{n\perp} \tilde{\mathcal{D}}_n \mathbf{v}_n - \tilde{\mathcal{E}}_{n\parallel} v_n \tilde{\mathcal{D}}_n = v_n \tilde{\mathcal{D}}_n \left[\tilde{\mathcal{E}}_{n\perp} \hat{\mathbf{k}}_n - \tilde{\mathcal{E}}_{n\parallel} \hat{\mathcal{E}}_{n\perp} \right] \end{aligned} \quad (C166)$$

若讲究对称也可写作：

$$\tilde{\mathbf{S}}_n = v_n \tilde{\mathcal{D}}_n \left[\operatorname{sgn}(\varepsilon_n^{(1)}) \tilde{\mathcal{E}}_{n\perp} \widehat{\mathcal{E}}_{n\parallel} - \tilde{\mathcal{E}}_{n\parallel} \widehat{\mathcal{E}}_{n\perp} \right]$$

$$\tilde{\mathbf{S}}_n = \tilde{\mathcal{E}}_{n\perp} \tilde{\mathcal{D}}_n \mathbf{v}_n - \tilde{\mathcal{E}}_{n\parallel} v_n \tilde{\mathcal{D}}_n$$

$$\tilde{\mathbf{S}}_n = (\tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n) \mathbf{v}_n - (\tilde{\mathcal{E}}_n \cdot \mathbf{v}_n) \tilde{\mathcal{D}}_n = \tilde{\mathcal{E}}_n \times (\mathbf{v}_n \times \tilde{\mathcal{D}}_n) = \tilde{\mathcal{E}}_n \times \tilde{\mathcal{H}}_n \quad (C167)$$

对比 $\tilde{\mathbf{S}}_n = \tilde{\mathbf{S}}_{n\parallel} + \tilde{\mathbf{S}}_{n\perp}$ 可得：

$$\tilde{\mathbf{S}}_{n\parallel} = (\tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n) \mathbf{v}_n$$

$$\tilde{\mathbf{S}}_{n\perp} = \tilde{\mathcal{S}}_{n\perp} \widehat{\mathcal{S}}_{n\perp} = -\tilde{\mathcal{E}}_{n\parallel} \tilde{\mathcal{H}}_n \widehat{\mathcal{D}}_n = -(\tilde{\mathcal{E}}_n \cdot \mathbf{v}_n) \tilde{\mathcal{D}}_n \quad (C168)$$

与之前的结论 $\tilde{\mathcal{S}}_{n\perp} = -\tilde{\mathcal{E}}_{n\parallel} \tilde{\mathcal{H}}_n$ 相照应。且对 $\tilde{\mathbf{S}}_n$ 点乘或叉乘 $\hat{\mathbf{k}}_n$ ，均可回到

$$\begin{aligned} \tilde{\mathcal{S}}_{n\parallel} &= \tilde{\mathcal{S}}_n \cos \alpha_n = \tilde{w}_n \frac{v_n}{\cos \alpha_n} \cos \alpha_n = \tilde{w}_n v_n \\ \tilde{\mathcal{S}}_{n\perp} &= \operatorname{sgn}(\varepsilon_n^{(1)}) \cdot \tilde{\mathcal{S}}_n \sin \alpha_n = \operatorname{sgn}(\varepsilon_n^{(1)}) \cdot \tilde{w}_n \frac{v_n}{\cos \alpha_n} \sin \alpha_n = \operatorname{sgn}(\varepsilon_n^{(1)}) \cdot \tilde{w}_n v_n \tan \alpha_n \\ &= \operatorname{sgn}(\varepsilon_n^{(1)}) \cdot \tilde{S}_{n\parallel} \tan \alpha_n \end{aligned} \quad (C169)$$

其中含有 $\tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n$ 即含有 $\cos \alpha_n$ 的 \tilde{w}_n 、 $\tilde{S}_{n\parallel}$ 与 $\tan \alpha_n$ 在 $\alpha_n > 0$ 时同号，在 $\alpha_n < 0$ 时异号，即 $\tilde{S}_{n\perp}$ 同号于 $\operatorname{sgn}(\varepsilon_n^{(1)}) \cdot \alpha_n$ ，符合物理情景，也说明该结论是正确的。

附C1.2.9.8 o光 e光的折射率曲面均分别来源于o光 e光各自不同极化方向而非偏振方向上的同一个折射率椭球所给出的不同极化程度

在线性介质中，

$$\tilde{w}_e = \frac{1}{2} \tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n = \frac{1}{2} [\tilde{\mathcal{E}}_{nx} \tilde{\mathcal{D}}_{nx} + \tilde{\mathcal{E}}_{ny} \tilde{\mathcal{D}}_{ny} + \tilde{\mathcal{E}}_{nz} \tilde{\mathcal{D}}_{nz}] = \frac{1}{2} \left[\frac{\tilde{\mathcal{D}}_{nx}^2}{\mathcal{E}_{nx}^{(1)}} + \frac{\tilde{\mathcal{D}}_{ny}^2}{\mathcal{E}_{ny}^{(1)}} + \frac{\tilde{\mathcal{D}}_{nz}^2}{\mathcal{E}_{nz}^{(1)}} \right] \quad (C170)$$

则有

$$\frac{(\tilde{\mathcal{D}}_{nx} / \sqrt{2\mathcal{E}_0 \tilde{w}_e})^2}{\mathcal{E}_{nx}^{(1)}} + \frac{(\tilde{\mathcal{D}}_{ny} / \sqrt{2\mathcal{E}_0 \tilde{w}_e})^2}{\mathcal{E}_{ny}^{(1)}} + \frac{(\tilde{\mathcal{D}}_{nz} / \sqrt{2\mathcal{E}_0 \tilde{w}_e})^2}{\mathcal{E}_{nz}^{(1)}} = \frac{\tilde{d}_{nx}^2}{n_{nex}^2} + \frac{\tilde{d}_{ny}^2}{n_{ney}^2} + \frac{\tilde{d}_{nz}^2}{n_{nez}^2} = 1 \quad (C171)$$

这便是 $\varepsilon_n^{(1)}$ 三分量均 > 0 条件下，光学意义上的折射率椭球，也称为 (n_n, \tilde{d}_n) 曲面，不过由于含有能量密度，以及相关物理量必须是实的，该式子几乎无用武之地，没有之前定义的折射率椭球实用，像极了爱因斯坦所认为的华而无实的数学技巧。

(1) 两个折射率曲面均来源于同一折射率椭球，是因波矢方向上的等效标量折射率来自且等价于极化方向的等效标量折射率

但它进一步说明折射率椭球考察的是极化(或磁化)方向上的等效标量折射率大小，而不是电矢量的偏振方向上的，现可对此进行更严谨的论证：对 o光已经做了折射率椭球与折射率曲面的转换的讨论，但无法严格得出折射率曲面在相速度方向的折射率，是来源于折射率椭球在偏振方向的折射率，还是极化方向的折射率；并且即使偏振方向与波矢方向的相对位置会随着波矢方向变化，二者仍然是一一对应关系，因此从这两个方面均无法否定考虑偏振方向的可能。

现将 $D_{nex} = 0$ 代入折射率椭球有

$$\begin{aligned} \frac{(n_{ne}x_0)^2}{n_{nex}^2} + \frac{(n_{ne}y_0)^2}{n_{ney}^2} + \frac{(n_{ne}z_0)^2}{n_{nez}^2} &= \frac{(n_{nee}D_{nex}/D_{ne})^2}{n_{ne\perp}^2} + \frac{(n_{nee}D_{ney}/D_{ne})^2}{n_{ne\perp}^2} + \frac{(n_{nee}D_{nez}/D_{ne})^2}{n_{ne//}^2} \\ &= \frac{(n_{nee}D_{ney}/D_{ne})^2}{n_{ne\perp}^2} + \frac{(n_{nee}D_{nez}/D_{ne})^2}{n_{ne//}^2} = \frac{[n_{nee} \sin(\theta_{ne} + \frac{\pi}{2})]^2}{n_{ne\perp}^2} + \frac{[n_{nee} \cos(\theta_{ne} + \frac{\pi}{2})]^2}{n_{ne//}^2} = 1 \quad (C172) \end{aligned}$$

对比 e光的折射率曲面

$$\frac{(n_{nee} \sin \theta_{ne})^2}{n_{ne//}^2} + \frac{(n_{nee} \cos \theta_{ne})^2}{n_{ne\perp}^2} = 1 \quad (C173)$$

可见二者本质上是相同的。同理就连可能的极化双曲面

$$\frac{\varepsilon_{ne}^{(1)} \sin^2(\theta_{ne} + \frac{\pi}{2})}{\varepsilon_{n\perp}^{(1)}} + \frac{\varepsilon_{ne}^{(1)} \cos^2(\theta_{ne} + \frac{\pi}{2})}{\varepsilon_{n\parallel}^{(1)}} = 1 \quad (C174)$$

与波速双曲面

$$\frac{\varepsilon_{ne}^{(1)} \sin^2 \theta_{ne}}{\varepsilon_{n\parallel}^{(1)}} + \frac{\varepsilon_{ne}^{(1)} \cos^2 \theta_{ne}}{\varepsilon_{n\perp}^{(1)}} = 1 \quad (C175)$$

也是等价的。因此 e 光的相速度方向上的折射率，本质上也来源于其极化方向上的折射率，且有

$$\frac{1}{\mu_n^{(1)} \varepsilon_{ne}^{(1)}} = \frac{\omega_n^2}{k_{ne}^2} = v_{ne}^2 \quad (C176)$$

于是 e 光也符合磁各向同性介质中的波动方程。并且 o 光和 e 光均同享一个旋转椭球面折射率椭球，只因 e 光和 o 光各自的偏振和极化方向的独特取向，导致二者的折射率曲面一个是球面一个是椭球。

(2) 每个波动方程的解仍只有一个，但两种极化方向导致两种等效标量介电常数，以至于出现两个波动方程，则有两个解

在这样的理解下，每个线性介质中齐次无源(无非线性波源)波动方程的解仍只有一个，但因折射率曲面及与折射率曲面相绑定的偏振方向有两个，则波动方程中的等效标量介电常数会随不同解的传播和偏振方向的不同而不同，而波动方程中介电常数的不同又会导致解的不同，即构成了方程-解-方程的耦合。而因 Snell 定律对波矢的约束，而可能存在两个等效标量介电常数值、两个波矢和相应的两种波动方程，所以总的解可能有两个，而由于这两个解的波矢必须符合 Snell 定律，因此两个解的波矢方向一般会有所差别，而极化方向一定会有差别，这就是双折射的本质。

(3) 偏振方向不同→极化方向不同→等效标量介电常数不同→波动方程不同，若偏振方向上等效标量介电常数为复数，则相应波动方程的解含衰减因子，导致二向色性

自然界中存在某些各向异性晶体，其在某个方向的等效标量介电常数是复数，因此若 o 光或 e 光的极化方向，与该方向相同，则代入其波动方程后所得的解会含 e 的负实指数因子，相应的偏振光强度会随距离迅速衰减，或者说被介质强烈地吸收，因此 o 光和 e 光只有垂直于该方向的分量的幅度才不会衰减，此即二向色性。

可利用这样的各向异性材料（如液晶）制造人造偏振片，以产生极化和振动方向恒定的线偏振光。所有各向异性晶体都可产生双折射现象，但不是所有各向异性晶体

都有二向色性（有些单轴晶体没有，有些双轴晶体有三向色性），所以 o、e 光的偏振方向的不同，与选择性吸收和选择性透过的方向的不同，是两个不同的概念。利用偏振片吸收掉某一束光，可以选择出射 o 光或 e 光，而利用波矢的分离、全反射，所制作的分束棱镜，也可实现这种选择，但其同时也保留了另一束^①。

由于一般介质除了少数是吸收介质外，还多是色散介质，吸收+色散的介质中，介电常数各分量的虚部和模的大小也与频率有关，因此波动方程及其解也与频率有关，因此晶体对光波的吸收既取决于入射光的传播和偏振方向，又取决于光的波长，此时对于复色入射光，晶体没有一个特定的吸收方向，导致折射光的强度、颜色都有一定的球面度分布，这称为多向色性。

附C1.2.9.9 双负材料与双正材料中波矢与能流方向的确定

(1) 波矢方向遵循折射定律、反射定律，但光线方向不遵循，需先确定波矢方向再通过离散角确定光线方向，或直接用菲涅尔光线方程即惠更斯作图法确定大小和方向

由于折射定律与反射定律是由电磁场的边值关系得到的，其中的相位匹配条件只适用于波矢，因此只有波矢方向遵循折射定律和反射定律，而折射光和反射光的光线方向，不遵循折射定律和反射定律。但只要入射面与主截面平行，不论是否斜入射，则 e 光和 o 光的光线方向仍均在入射面内：因为此时 o 光和 e 光的波法线方向均在入射面内，虽然 o 光的偏振和极化方向垂直于自己的主平面即主截面或入射面，但 o 光的光线方向与 o 光波法线方向相同，并没有朝着或反向于光轴方向偏离；同时 e 光的偏振和极化方向也在入射面内，则虽然 e 光的光线方向也会偏向或偏离光轴方向，但 e 光光线方向与其 $\mathbf{k}_{ne}, \mathbf{E}_{ne}(\mathbf{r}), \mathbf{D}_{ne}(\mathbf{r})$ 共平面，所以也在入射面内，因此当 $\mathbf{k}_{ne}, \mathbf{k}_{no}, \mathbf{k}$ 三者共面时，必同时有 $\mathbf{S}_{ne}, \mathbf{S}_{no}, \mathbf{k}_{ne}, \mathbf{k}_{no}, \mathbf{k}$ 五者共面，且与入射光、界面法线、e 光的极化和偏振方向一起共面。但当 $\mathbf{k}_{ne}, \mathbf{k}_{no}, \mathbf{k}$ 三者不共面时，一般 $\mathbf{S}_{ne}, \mathbf{S}_{no}, \mathbf{k}$ 也就不共面了。具体来说， $\mathbf{S}_{ne}, \mathbf{S}_{no}$ 的方向可通过惠更斯作图法确定，或先确定波法线方向，再由离散角公式计算光线方向。

^① 双折射本身是用于获得两束偏振光，并自带同时分离两束光的波矢和相位。而外加其他的特性如复介电常数分量可以吸收掉某束光（偏振片）、外加角度的巧妙设置可实现全反射分离两束光（分光棱镜）、外加厚度的精确不连续设置可准确转换/切换光的几种偏振状态（波片）、外加厚度的可连续调谐则可连续地改变双光束的相位差（补偿器）。

(2) 双负材料中 $\mathbf{E}_n, \mathbf{H}_n, \mathbf{k}_n$ 变为左手系, 场能密度为负, 能流密度反向

目前理论上已证明自然界不存在天然的负折射率材料, 但现在通过人工微结构已经制造出具有负折射率的材料, 其介电常数 $\vec{\epsilon}^{(1)}$ 和磁导率 $\vec{\mu}^{(1)}$ 均为负 (但一般只在某个方向上的等效标量 $\mu_n^{(1)}, \epsilon_n^{(1)}$ 为负, 或张量的某些分量为负), 称为双负材料或左手材料: 假设两材料除了折射率相反外, 其他条件一样, 且假设其中折射光 (不论是 e 光还是 o 光还是其他, 但得都是同类) 的 $\mathbf{E}_n, \mathbf{k}_n$ 相同, 则两种材料中折射光的 $\mathbf{E}_n, \mathbf{B}_n, \mathbf{k}_n, \mathbf{v}_n$ 相同, $\mathbf{D}_n, \mathbf{H}_n, \mathbf{S}_n, \mathbf{v}_{nr}$ 均反向, 正常材料中的右手系 $\mathbf{E}_n, \mathbf{H}_n, \mathbf{k}_n$ 、 $\mathbf{D}_n, \mathbf{B}_n, \mathbf{k}_n$ 在双负材料中均变为左手系, 所以双负材料又叫左手材料。正常材料中的右手系 $\mathbf{E}_n, \mathbf{B}_n, \mathbf{k}_n$ 、 $\mathbf{D}_n, \mathbf{H}_n, \mathbf{k}_n$ 、 $\mathbf{E}_n, \mathbf{H}_n, \mathbf{S}_n$ 在人工双负材料中仍保持右手系不变。

与此同时, 双负材料中的能量密度 $\tilde{w}_n = 2\tilde{w}_{ne} = \tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n = \tilde{\mathcal{E}}_{n\perp} \tilde{\mathcal{D}}_n = \epsilon_n^{(1)} \tilde{\mathcal{E}}_{n\perp}^2 < 0$ 与正常介质中的反号, 为负, 与能流方向反向一起, 表示能量并非从外界入射光射入晶体, 反而从晶体向外界环境输出能量, 此时单色平面波因能流 $\mathbf{S}_n, \mathbf{v}_{nr}$ 大致反平行于波法线 $\mathbf{k}_n, \mathbf{v}_n$ 而看上去是“向后传播的”。但一般对于相同入射光, 双负介质和双正介质中折射光的 $\mathbf{E}_n, \mathbf{k}_n$ 一般不同, 所以这样的现象需要对两种介质构造不同的入射光和晶面的夹角和入射光偏振才可能发生。

(3) 左手材料与右手材料均遵循电磁场边值关系、满足同样的 Snell 定律, 但无法得出左手材料中的波矢负折射结论

考察两束完全相同的入射光入射这样的双负介质 (设也是线性的) 和另一个除折射率为负外, 其他参数相同的理想的正常(线性)介质, 则根据电磁场的边值关系中与含时麦氏方程组三号方程对应的方程, 电矢量于平行于分界面方向的切向分量连续, 即有位于界面之上同侧的入射光和反射光的电矢量和的切向分量, 与界面下晶体中折射光的电矢量切向分量满足

$$\tilde{\mathbf{E}}_{nT} \Big|_{S^+} + \tilde{\mathbf{E}}'_{nT} \Big|_{S^+} = \tilde{\mathbf{E}}''_{nT} \Big|_{S^-} \quad (C177)$$

同理有

$$\begin{aligned} \tilde{\mathbf{H}}_{nT} \Big|_{S^+} + \tilde{\mathbf{H}}'_{nT} \Big|_{S^+} &= \tilde{\mathbf{H}}''_{nT} \Big|_{S^-} \\ \tilde{\mathbf{D}}_{nN} \Big|_{S^+} + \tilde{\mathbf{D}}'_{nN} \Big|_{S^+} &= \tilde{\mathbf{D}}''_{nN} \Big|_{S^-} \\ \tilde{\mathbf{B}}_{nN} \Big|_{S^+} + \tilde{\mathbf{B}}'_{nN} \Big|_{S^+} &= \tilde{\mathbf{B}}''_{nN} \Big|_{S^-} \end{aligned} \quad (C178)$$

含时物理量的切向和法向连续意味着定振幅随介质连续的同时, 相位部分也需要

统一，即相位中的空间部分与时间部分也需分别相等（比连续高一级，因为入射和反射波的相位也得匹配，否则即使定振幅连续，含时物理量也不连续），则有：

$$\begin{aligned} \mathbf{E}_{0nT} \Big|_{S^+} + \mathbf{E}'_{0nT} \Big|_{S^+} &= \mathbf{E}''_{0nT} \Big|_{S^-} \\ \mathbf{H}_{0nT} \Big|_{S^+} + \mathbf{H}'_{0nT} \Big|_{S^+} &= \mathbf{H}''_{0nT} \Big|_{S^-} \\ \mathbf{D}_{0nN} \Big|_{S^+} + \mathbf{D}'_{0nN} \Big|_{S^+} &= \mathbf{D}''_{0nN} \Big|_{S^-} \\ \mathbf{B}_{0nN} \Big|_{S^+} + \mathbf{B}'_{0nN} \Big|_{S^+} &= \mathbf{B}''_{0nN} \Big|_{S^-} \\ \mathbf{k}_n \cdot \mathbf{r} \Big|_{S^+} = \mathbf{k}'_n \cdot \mathbf{r} \Big|_{S^+} &= \mathbf{k}''_n \cdot \mathbf{r} \Big|_{S^-} \\ \omega_n t \Big|_{S^+} = \omega'_n t \Big|_{S^+} &= \omega''_n t \Big|_{S^-} \end{aligned} \tag{C179}$$

其中三个波的(\mathbf{r}, t)均需采用同一套时空四维坐标系，才能配合上述关系实现相位一致。现采用以晶面为 $x-o-y$ 面、 z 轴垂直于晶面指向介质外的直角坐标系，则

$$k_{nx}x + k_{ny}y + k_{nz}z \Big|_{z=0^+} = k'_{nx}x + k'_{ny}y + k'_{nz}z \Big|_{z=0^+} = k''_{nx}x + k''_{ny}y + k''_{nz}z \Big|_{z=0^-} \tag{C180}$$

由于入射前、折射和反射后的波矢和频率与入射折射反射发生时的保持不变，有

$$k_{nx}x + k_{ny}y = k'_{nx}x + k'_{ny}y = k''_{nx}x + k''_{ny}y \quad \omega_n = \omega'_n = \omega''_n \tag{C181}$$

其中波矢部分可改写作 $\mathbf{k}_{nT} = \mathbf{k}'_{nT} = \mathbf{k}''_{nT}$ ，或 $\mathbf{n} \times \mathbf{k}_n = \mathbf{n} \times \mathbf{k}'_n = \mathbf{n} \times \mathbf{k}''_n$ ，进而取模有

$$\begin{aligned} k_n \sin \langle \mathbf{k}_n, \mathbf{n} \rangle &= k'_n \sin \langle \mathbf{k}'_n, \mathbf{n} \rangle = k''_n \sin \langle \mathbf{k}''_n, \mathbf{n} \rangle, \text{ 即 } k_n \sin \gamma_n = k'_n \sin \gamma'_n = k''_n \sin \gamma''_n, \text{ 即} \\ \frac{\omega_n |n_n|}{c} \sin \gamma_n &= \frac{\omega'_n |n'_n|}{c} \sin \gamma'_n = \frac{\omega''_n |n''_n|}{c} \sin \gamma''_n \end{aligned} \tag{C182}$$

又因 $\omega_n = \omega'_n = \omega''_n$ ，于是有 $|n_n| \sin \gamma_n = |n'_n| \sin \gamma'_n = |n''_n| \sin \gamma''_n$ ，其中描述入射与折射关系的部分即 Snell 定律。

从上述过程可看出，相位一致并没法推导出负折射率材料中相应折射角 $\gamma''_n = \langle \mathbf{k}''_n, \mathbf{n} \rangle$ 反号并为负的结论，何况它没法为负。因此许多文献中的结论不成立，只能说通过去掉折射率的绝对值符号，以巧妙地构造出了类似的方程，恰好确实解释了真实存在的客观现象。

(4) 考察左右手材料边界上电磁场分量的异同，导出左手材料中能流的负折射，波矢并非负折射

①. 两种材料内部靠近边界处，及其内部空间中电磁场分量的异同

真实结论比不仅与它有别，得来过程也不同：假定入射左手和右手材料的两束入

射光的各参数均为 $E_{0n}, H_{0n}, D_{0n}, B_{0n}, k_n, \omega_n, \gamma_n$ ，同时也假设从左手和右手材料反射的两束反射光的光学参数相同 $E'_{0n}, H'_{0n}, D'_{0n}, B'_{0n}, k'_n, \omega'_n, \gamma'_n$ ，则据电磁波的边值关系的约束，左手材料和右手材料下表面的 $E''_{0nT} \Big|_{S^-}, H''_{0nT} \Big|_{S^-}, D''_{0nN} \Big|_{S^-}, B''_{0nN} \Big|_{S^-}, k''_{nT}, \omega''_n, \gamma''_n$ 相同。

但既然两种材料下表面的 $D''_{0nN} \Big|_{S^-} = \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot E''_{0nN} \Big|_{S^-}, B''_{0nN} \Big|_{S^-} = \overset{\rightarrow}{\mu}_n^{(1)} \cdot H''_{0nN} \Big|_{S^-}$ 相同，而两种材料的 $\overset{\rightarrow}{\epsilon}_n^{(1)}, \overset{\rightarrow}{\mu}_n^{(1)}$ 反号，因此两种材料下表面的 $E''_{0nN} \Big|_{S^-}, H''_{0nN} \Big|_{S^-}$ 反号；同理，两种材料下表面的 $E''_{0nT} \Big|_{S^-}, H''_{0nT} \Big|_{S^-}$ 相同、 $\overset{\rightarrow}{\epsilon}_n^{(1)}, \overset{\rightarrow}{\mu}_n^{(1)}$ 反号，则二者下表面的 $D''_{0nT} \Big|_{S^-} = \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot E''_{0nT} \Big|_{S^-}, B''_{0nT} \Big|_{S^-} = \overset{\rightarrow}{\mu}_n^{(1)} \cdot H''_{0nT} \Big|_{S^-}$ 反号。

那么若记右手材料下表面的剩余四个场参数为 $E''_{0nN} \Big|_{S^-}, H''_{0nN} \Big|_{S^-}, D''_{0nT} \Big|_{S^-}, B''_{0nT} \Big|_{S^-}$ ，则左手材料下表面的剩余四项参数为 $-E''_{0nN} \Big|_{S^-}, -H''_{0nN} \Big|_{S^-}, -D''_{0nT} \Big|_{S^-}, -B''_{0nT} \Big|_{S^-}$ 。若将上述两种材料下表面各参数的 $|_{S^-}$ 条件去掉，上述关系式仍成立，拓展到含时含空的瞬时值也成立，再拓展到实数形式也成立。

②. 双负材料内部折射光光线与入射光光线在法线同侧，即法向分量与双正材料相同、切向分量与双正材料相反，二者及二者的方向均关于法线对称

那么左手材料中折射光的能流密度矢量

$$\begin{aligned} \tilde{S}_{n \leftarrow}'' &= \tilde{\mathcal{E}}_{n \leftarrow}'' \times \tilde{\mathcal{H}}_{n \leftarrow}'' = (\tilde{\mathcal{E}}_{n \leftarrow N}'' + \tilde{\mathcal{E}}_{n \leftarrow T}'') \times (\tilde{\mathcal{H}}_{n \leftarrow N}'' + \tilde{\mathcal{H}}_{n \leftarrow T}'') = \tilde{S}_{n \leftarrow N}'' + \tilde{S}_{n \leftarrow T}'' \\ &= (\tilde{\mathcal{E}}_{n \leftarrow N}'' \times \tilde{\mathcal{H}}_{n \leftarrow N}'' + \tilde{\mathcal{E}}_{n \leftarrow T}'' \times \tilde{\mathcal{H}}_{n \leftarrow T}'') + (\tilde{\mathcal{E}}_{n \leftarrow N}'' \times \tilde{\mathcal{H}}_{n \leftarrow T}'' + \tilde{\mathcal{E}}_{n \leftarrow T}'' \times \tilde{\mathcal{H}}_{n \leftarrow N}'') \\ &= (\tilde{\mathcal{E}}_{n \rightarrow N}'' \times \tilde{\mathcal{H}}_{n \rightarrow N}'' + \tilde{\mathcal{E}}_{n \rightarrow T}'' \times \tilde{\mathcal{H}}_{n \rightarrow T}'') - (\tilde{\mathcal{E}}_{n \rightarrow N}'' \times \tilde{\mathcal{H}}_{n \rightarrow T}'' + \tilde{\mathcal{E}}_{n \rightarrow T}'' \times \tilde{\mathcal{H}}_{n \rightarrow N}'') \\ &= \tilde{S}_{n \rightarrow N}'' - \tilde{S}_{n \rightarrow T}'' = (\tilde{S}_{n \rightarrow N}'' + \tilde{S}_{n \rightarrow T}'') - 2\tilde{S}_{n \rightarrow T}'' = \tilde{S}_{n \rightarrow}'' - 2\tilde{S}_{n \rightarrow T}'' \coloneqq \tilde{S}_n'' - 2\tilde{S}_{nT}'' \end{aligned} \quad (C183)$$

从中可见

$$\tilde{S}_{n \leftarrow N}'' = \tilde{S}_{n \rightarrow N}'' \quad \text{and} \quad \tilde{S}_{n \leftarrow T}'' = -\tilde{S}_{n \rightarrow T}'' \quad (C184)$$

因此还有

$$\tilde{S}_n'' = \tilde{S}_{n \rightarrow}'' = \tilde{S}_{n \rightarrow N}'' + \tilde{S}_{n \rightarrow T}'' = \tilde{S}_{n \leftarrow N}'' - \tilde{S}_{n \leftarrow T}'' = \tilde{S}_{n \leftarrow}'' - 2\tilde{S}_{n \leftarrow T}'' \quad (C185)$$

可见双负材料中入射光光线与折射光光线在法线同侧。

③. 双负材料内部折射波矢虽与折射光线同象限/卦限，但其法向分量与双正材料相反，切向分量与双正材料相同，二者方向关于界面对称

考虑左手介质中与右手介质中的折射光的 $\tilde{\mathcal{E}}_{n \rightarrow}, k_n$ 相同的光，之前提到过，其 $\tilde{\mathcal{E}}_n, \tilde{\mathcal{B}}_n, k_n, v_n$ 与右手介质相同， $\tilde{\mathcal{D}}_n, \tilde{\mathcal{H}}_n, S_n, v_{nr}$ 均反向，因此 $\tilde{S}_{n \leftarrow}'' = -\tilde{S}_{n \rightarrow}''$ 、 $\tilde{\mathcal{D}}_{n \leftarrow}'' = -\tilde{\mathcal{D}}_{n \rightarrow}''$ 、

$\tilde{\mathcal{E}}_{n \leftarrow}'' = \tilde{\mathcal{E}}_{n \rightarrow}''$ ，那么据(C184)有

$$\tilde{\mathbf{S}}_{n \leftarrow N}'' = \tilde{\mathbf{S}}_{n \rightarrow N}'' = -\tilde{\mathbf{S}}_{n \leftarrow N}'' \quad , \quad \tilde{\mathbf{S}}_{n \leftarrow T}'' = -\tilde{\mathbf{S}}_{n \rightarrow T}'' = \tilde{\mathbf{S}}_{n \leftarrow T}'' \quad (C186)$$

同理，据电位移矢量在左右手材料中的法向切向边值关系 $\tilde{\mathcal{D}}_{n \leftarrow N}'' = \tilde{\mathcal{D}}_{n \rightarrow N}''$ 、
 $\tilde{\mathcal{D}}_{n \leftarrow T}'' = -\tilde{\mathcal{D}}_{n \rightarrow T}''$ ，也有

$$\tilde{\mathcal{D}}_{n \leftarrow N}'' = \tilde{\mathcal{D}}_{n \rightarrow N}'' = -\tilde{\mathcal{D}}_{n \leftarrow N}'' \quad , \quad \tilde{\mathcal{D}}_{n \leftarrow T}'' = -\tilde{\mathcal{D}}_{n \rightarrow T}'' = \tilde{\mathcal{D}}_{n \leftarrow T}'' \quad (C187)$$

而据 $\tilde{\mathcal{E}}_{n \leftarrow T}'' = \tilde{\mathcal{E}}_{n \rightarrow T}''$ 、 $\tilde{\mathcal{E}}_{n \leftarrow N}'' = -\tilde{\mathcal{E}}_{n \rightarrow N}''$ ，也有

$$\tilde{\mathcal{E}}_{n \leftarrow N}'' = -\tilde{\mathcal{E}}_{n \rightarrow N}'' = -\tilde{\mathcal{E}}_{n \leftarrow N}'' \quad , \quad \tilde{\mathcal{E}}_{n \leftarrow T}'' = \tilde{\mathcal{E}}_{n \rightarrow T}'' = \tilde{\mathcal{E}}_{n \leftarrow T}'' \quad (C188)$$

可见左手介质中的真实折射光的 $\tilde{\mathcal{E}}_{n \leftarrow}''$ 、 $\tilde{\mathbf{S}}_{n \leftarrow}''$ 、 $\tilde{\mathcal{D}}_{n \leftarrow}''$ ，与左手介质中与右手介质折射光波矢和电矢量重合的假想折射光的 $\tilde{\mathcal{E}}_{n \leftarrow}''$ 、 $\tilde{\mathbf{S}}_{n \leftarrow}''$ 、 $\tilde{\mathcal{D}}_{n \leftarrow}''$ ，关于分界面对称。

而由于任何坐标系、任何线性介质下均应有 $\langle \tilde{\mathbf{S}}_n, \mathbf{k}_n \rangle = \langle \tilde{\mathcal{E}}_n, \tilde{\mathcal{D}}_n \rangle$ ，那么当右手介质中的折射光满足该结论 $\langle \tilde{\mathbf{S}}_{n \rightarrow}'', \mathbf{k}_{n \rightarrow}'' \rangle = \langle \tilde{\mathcal{E}}_{n \rightarrow}'', \tilde{\mathcal{D}}_{n \rightarrow}'' \rangle$ 时，之前也已经给出与右手介质折射光的 $\tilde{\mathcal{E}}_{n \rightarrow}''$ 、 $\mathbf{k}_{n \rightarrow}''$ 相同的左手介质中的假想折射光也满足 $\langle \tilde{\mathbf{S}}_{n \leftarrow}'', \mathbf{k}_{n \leftarrow}'' \rangle = \langle \tilde{\mathcal{E}}_{n \leftarrow}'', \tilde{\mathcal{D}}_{n \leftarrow}'' \rangle$ ，那么要想左手介质中真正的折射光也满足 $\langle \tilde{\mathbf{S}}_{n \leftarrow}'', \mathbf{k}_{n \leftarrow}'' \rangle = \langle \tilde{\mathcal{E}}_{n \leftarrow}'', \tilde{\mathcal{D}}_{n \leftarrow}'' \rangle$ ，则由于现在已经有

$\tilde{\mathcal{E}}_{n \leftarrow}''$ 、 $\tilde{\mathbf{S}}_{n \leftarrow}''$ 、 $\tilde{\mathcal{D}}_{n \leftarrow}''$ 与 $\tilde{\mathcal{E}}_{n \leftarrow}''$ 、 $\tilde{\mathbf{S}}_{n \leftarrow}''$ 、 $\tilde{\mathcal{D}}_{n \leftarrow}''$ 关于交界面对称，那么必须同时满足 $\mathbf{k}_{n \leftarrow}''$ 与 $\mathbf{k}_{n \leftarrow}'' = \mathbf{k}_{n \rightarrow}''$ 关于交界面对称，左手介质中的真实折射光的能流与波矢的夹角，才等于电矢量与电位移矢量的夹角。

因此有 $\mathbf{k}_{n \leftarrow N}'' = -\mathbf{k}_{n \leftarrow N}''$ 、 $\mathbf{k}_{n \leftarrow T}'' = \mathbf{k}_{n \leftarrow T}''$ ，而因 $\mathbf{k}_{n \leftarrow}'' = \mathbf{k}_{n \rightarrow}''$ ，即有

$$\mathbf{k}_{n \leftarrow N}'' = -\mathbf{k}_{n \rightarrow N}'' \quad , \quad \mathbf{k}_{n \leftarrow T}'' = \mathbf{k}_{n \rightarrow T}'' \quad (C189)$$

若继续考察左手介质中真实和假想折射光的其他电磁场分量，则还有：

$$\begin{aligned} \tilde{\mathcal{B}}_{n \leftarrow N}'' &= \tilde{\mathcal{B}}_{n \rightarrow N}'' = \tilde{\mathcal{B}}_{n \leftarrow N}'' \quad , \quad \tilde{\mathcal{B}}_{n \leftarrow T}'' = -\tilde{\mathcal{B}}_{n \rightarrow T}'' = -\tilde{\mathcal{B}}_{n \leftarrow T}'' \\ \tilde{\mathcal{H}}_{n \leftarrow N}'' &= -\tilde{\mathcal{H}}_{n \rightarrow N}'' = \tilde{\mathcal{H}}_{n \leftarrow N}'' \quad , \quad \tilde{\mathcal{H}}_{n \leftarrow T}'' = \tilde{\mathcal{H}}_{n \rightarrow T}'' = -\tilde{\mathcal{H}}_{n \leftarrow T}'' \end{aligned} \quad (C190)$$

意味着左手介质中的真实折射光除了可以看做由右手介质中的真正折射光通过将其 $\tilde{\mathcal{E}}_{n \rightarrow}''$ 、 $\tilde{\mathcal{H}}_{n \rightarrow}''$ 、 $\mathbf{k}_{n \rightarrow}''$ 沿界面翻折的同时，将其 $\tilde{\mathcal{D}}_{n \rightarrow}''$ 、 $\tilde{\mathcal{B}}_{n \rightarrow T}''$ 、 $\tilde{\mathbf{S}}_{n \rightarrow}''$ 沿法线翻折得到；也可以看做将右手介质中假想负折射光的 $\tilde{\mathcal{E}}_{n \leftarrow}''$ 、 $\tilde{\mathcal{S}}_{n \leftarrow}''$ 、 $\tilde{\mathcal{D}}_{n \leftarrow}''$ 、 $\mathbf{k}_{n \leftarrow}''$ 沿界面面对称的同时， $\tilde{\mathcal{B}}_{n \leftarrow}''$ 、 $\tilde{\mathcal{H}}_{n \leftarrow}''$ 沿法线轴对称来得到。——并且这后一种得来方法启示我们将 $\mathbf{i}_{n \rightarrow}$ 、 $\mathbf{j}_{n \rightarrow}$ 、 \mathbf{k} ，也与在 $z-o-y$ (\mathbf{k}, \mathbf{j}_n 平面) 平面上的 $\tilde{\mathcal{E}}_{n \leftarrow}''$ 、 $\tilde{\mathbf{S}}_{n \leftarrow}''$ 、 $\tilde{\mathcal{D}}_{n \leftarrow}''$ 、 $\mathbf{k}_{n \leftarrow}''$ 一起沿界面对称后，作为左手介质中真实折射光的坐标系 $\mathbf{i}_{n \leftarrow}$ 、 $\mathbf{j}_{n \leftarrow}$ 、 \mathbf{k} 。在这样的坐标系下， $\tilde{\mathcal{E}}_{n \leftarrow}''$ 、 $\tilde{\mathbf{S}}_{n \leftarrow}''$ 、 $\tilde{\mathcal{D}}_{n \leftarrow}''$ 、 $\mathbf{k}_{n \leftarrow}''$ 四个量相对于

此坐标系的角度可用旧有公式描述，与相应的 $\tilde{\mathcal{E}}_{n\leftarrow}''$, $\tilde{\mathbf{S}}_{n\leftarrow}''$, $\tilde{\mathcal{D}}_{n\leftarrow}''$, $\mathbf{k}_{n\leftarrow}''$ 相对于原 $\mathbf{i}_{n\leftarrow}$, $\mathbf{j}_{n\leftarrow}$, \mathbf{k} 即原 $\mathbf{i}_{n\rightarrow}$, $\mathbf{j}_{n\rightarrow}$, \mathbf{k} 的角度相同。

可见，只有先确定左手介质折射光的 $\tilde{\mathbf{S}}_{n\leftarrow}''$ 后，即在确定了 $\tilde{\mathcal{E}}_{n\leftarrow}''$, $\tilde{\mathcal{D}}_{n\leftarrow}''$, $\tilde{\mathbf{S}}_{n\leftarrow}''$ 后，才可以确定左手介质折射光的 $\mathbf{k}_{n\leftarrow}''$ ，因此从这个角度，能流密度比波矢更基本。

④. 双正双负材料中的折射光波矢沿界面对称的标量描述

波矢方向由能流方向确定的另一种标量描述：只有在确定了左手材料中折射光能流密度 $\tilde{\mathbf{S}}_{n\leftarrow}''$ 相对于界面外法线的方向即 $\langle \tilde{\mathbf{S}}_{n\leftarrow}'', \mathbf{n} \rangle$ 后，才能根据的离散角公式（其中 $\theta_{n\leftarrow}'' = \langle \mathbf{k}_{n\leftarrow}'', \mathbf{k} \rangle$ ）

$$\langle \tilde{\mathbf{S}}_{n\leftarrow}'', \mathbf{k}_{n\leftarrow}'' \rangle = \alpha_{n\leftarrow}'' = \arctan \left[\frac{1}{2} \sin 2\theta_{n\leftarrow}'' \left(\frac{1}{\varepsilon_{n\leftarrow/\parallel}^{(1)}} - \frac{1}{\varepsilon_{n\leftarrow/\perp}^{(1)}} \right) \cdot \varepsilon_{n\leftarrow}^{(1)} \right] + \pi \quad (C191)$$

确定此时折射光波法线 $\mathbf{k}_{n\leftarrow}''$ 相对于界面外法线的方向。

但在此之前，首先需要对 $\langle \tilde{\mathbf{S}}_{n\leftarrow}'', \mathbf{n} \rangle$ 、 $\langle \tilde{\mathbf{S}}_{n\rightarrow}'', \mathbf{n} \rangle$ 、 $\langle \mathbf{k}_n, \mathbf{n} \rangle$ 、 $\langle \mathbf{k}_n', \mathbf{n} \rangle$ 、 $\langle \mathbf{k}_n'', \mathbf{n} \rangle$ 等角度选定同一套坐标系，并规定相应物理量相对于 \mathbf{n} 的正向旋转方向为 $\mathbf{i}_{n\rightarrow}$, $\mathbf{j}_{n\rightarrow}$, \mathbf{k} 中 \mathbf{k} 到 $\mathbf{j}_{n\rightarrow}$ 的锐角方向。在这样以 \mathbf{n} 为 x 轴 or 极轴、以 \mathbf{k} 到 $\mathbf{j}_{n\rightarrow}$ 为正向的坐标系下，会发现 Snell 定律 $|n_n| \sin \gamma_n = |n_{n\leftarrow}| \sin \gamma_{n\leftarrow}'' = |n_{n\rightarrow}| \sin \gamma_{n\rightarrow}''$ 仍成立，并且为此左手介质折射光波矢 $\mathbf{k}_{n\leftarrow}''$ 方向必须沿是界面而不是沿法线翻折；另外在此坐标系下，式(C184)可写作

$$\langle \tilde{\mathbf{S}}_{n\leftarrow}'', \mathbf{n} \rangle = -\langle \tilde{\mathbf{S}}_{n\rightarrow}'', \mathbf{n} \rangle.$$

前节已证明式(C191)在 $\mathbf{i}_{n\leftarrow}$, $\mathbf{j}_{n\leftarrow}$, \mathbf{k} 系下仍成立以描述和度量左手介质中的折射光各矢量的角度，但正因如此，它所度量出的角度，是相对于该系的角度。而该 $\mathbf{i}_{n\leftarrow}$, $\mathbf{j}_{n\leftarrow}$, \mathbf{k} 系是通过 $\mathbf{i}_{n\rightarrow}$, $\mathbf{j}_{n\rightarrow}$, \mathbf{k} 沿界面翻折所得，因此两个坐标系的角度旋转正向方向相反，因而所度量得到的相同物理量在各自系下的角度，也相反。于是在 $\mathbf{i}_{n\leftarrow}$, $\mathbf{j}_{n\leftarrow}$, \mathbf{k} 系的量度下：

$$\gamma_{n\leftarrow}'' \Big|_{\leftarrow} = \langle \mathbf{k}_{n\leftarrow}'', \mathbf{n} \rangle \Big|_{\leftarrow} = \langle \mathbf{k}_{n\leftarrow}'', \tilde{\mathbf{S}}_{n\leftarrow}'' \rangle \Big|_{\leftarrow} - \langle \mathbf{n}, \tilde{\mathbf{S}}_{n\leftarrow}'' \rangle \Big|_{\leftarrow} = -\alpha_{n\leftarrow}'' \Big|_{\leftarrow} + \langle \tilde{\mathbf{S}}_{n\leftarrow}'', \mathbf{n} \rangle \Big|_{\leftarrow} \quad (C192)$$

那么在 $\mathbf{i}_{n\rightarrow}$, $\mathbf{j}_{n\rightarrow}$, \mathbf{k} 系的量度下，有：

$$\gamma_{n\leftarrow}'' \Big|_{\rightarrow} = -\alpha_{n\leftarrow}'' \Big|_{\rightarrow} + \langle \tilde{\mathbf{S}}_{n\leftarrow}'', \mathbf{n} \rangle \Big|_{\rightarrow} = \alpha_{n\leftarrow}'' \Big|_{\leftarrow} + \langle \tilde{\mathbf{S}}_{n\leftarrow}'', \mathbf{n} \rangle \Big|_{\rightarrow} = -\gamma_{n\leftarrow}'' \Big|_{\leftarrow} \quad (C193)$$

可以验证，由离散角以及 $\langle \tilde{\mathbf{S}}_{n\leftarrow}'', \mathbf{n} \rangle$ 所决定的该方向的 $\mathbf{k}_{n\leftarrow}''$ 符合 Snell 定律 $|n_n| \sin \gamma_n = |n_{n\leftarrow}| \sin \gamma_{n\leftarrow}''$ 这第二个对角度 $\gamma_{n\leftarrow}'' = \langle \mathbf{k}_{n\leftarrow}'', \mathbf{n} \rangle$ 的约束方程：

$$\begin{aligned} \text{根据 } |n_n| \sin \gamma_n &= |n_{n\leftarrow}| \sin \gamma_{n\leftarrow}'' = |n_{n\rightarrow}| \sin \gamma_{n\rightarrow}'' \text{ 以及 } |n_{n\leftarrow}| = |n_{n\rightarrow}|, \text{ 得} \\ \sin \gamma_{n\leftarrow}'' &= \sin \gamma_{n\rightarrow}'' \end{aligned} \quad (C194)$$

其中由于两旁均有同类物理量，因此不需要考虑是从哪个系量度的，因为若在一

个系下成立，则在另一个系下若二者均反号，等于方程两边同时取负，则也成立。

另一方面又因 $\langle \tilde{\mathbf{S}}''_{n\leftarrow}, \mathbf{n} \rangle = -\langle \tilde{\mathbf{S}}''_{n\rightarrow}, \mathbf{n} \rangle$ ，以及用双正介质中的真实折射光方向，描述双正介质中假想负折射光方向的，或通过双负介质中的假想正折射光方向，描述双负介质中的真实折射光（这里是后者）的 $\alpha''_n(-|\varepsilon_n^{(1)}|) = \alpha''_n(|\varepsilon_n^{(1)}|) + \pi$ ，即

$$\alpha''_{n\leftarrow}|_{\leftarrow} = \alpha''_{n\rightarrow}|_{\leftarrow} + \pi = -\alpha''_{n\rightarrow}|_{\rightarrow} + \pi = \alpha''_{n\rightarrow}|_{\rightarrow} + \pi \quad (C195)$$

将二者代入式(C193)得：

$$\begin{aligned} \gamma''_{n\leftarrow}|_{\rightarrow} &= \alpha''_{n\leftarrow}|_{\leftarrow} + \langle \tilde{\mathbf{S}}''_{n\leftarrow}, \mathbf{n} \rangle|_{\rightarrow} = \alpha''_{n\rightarrow}|_{\rightarrow} + \pi - \langle \tilde{\mathbf{S}}''_{n\rightarrow}, \mathbf{n} \rangle|_{\rightarrow} \\ &= \pi - \left(-\alpha''_{n\rightarrow}|_{\rightarrow} + \langle \tilde{\mathbf{S}}''_{n\rightarrow}, \mathbf{n} \rangle|_{\rightarrow} \right) = \pi - \gamma''_{n\rightarrow}|_{\rightarrow} \end{aligned} \quad (C196)$$

即有

$$\gamma''_{n\leftarrow} = \pi - \gamma''_{n\rightarrow} \quad (C197)$$

可见该结论符合式(C194) $\sin \gamma''_{n\leftarrow} = \sin \gamma''_{n\rightarrow}$ 。

虽然二者无矛盾，但 Snell 定律只能决定 $\mathbf{k}_{n\leftarrow}''$ 所在的直线，甚至只能决定可能存在 $\mathbf{k}_{n\leftarrow}''$ 的双锥面与入射面的两条交线，而 $\mathbf{k}_{n\leftarrow}''$ 方向在四条射线的方向的哪条射线方向上，需要通过先后计算 $\tilde{\mathbf{S}}''_{n\leftarrow}$ 的绝对位置和 α''_n 给出的 $\mathbf{k}_{n\leftarrow}''$ 相对于 $\tilde{\mathbf{S}}''_{n\leftarrow}$ 的位置，才能决定。此后，还可通过 $k_n \sin \theta = k_{n\leftarrow}'' \sin \theta_{\leftarrow}''$ 或 $k_{n\leftarrow}'' = \omega''_{n\leftarrow} |n_{n\leftarrow}''| / c$ 确定该方向的 $\mathbf{k}_{n\leftarrow}''$ 的大小 $k_{n\leftarrow}''$ 。

当然，确定大小的更优办法是：将式(C194)代入 $k_{n\leftarrow}'' \sin \gamma''_{n\leftarrow} = k_{n\rightarrow}'' \sin \gamma''_{n\rightarrow}$ ，或将 $|n_{n\leftarrow}''| = |n_{n\rightarrow}''|$ 、
 $\omega''_{n\leftarrow} = \omega''_{n\rightarrow}$ 代入 $k_{n\leftarrow}'' = \omega''_n |n_{n\leftarrow}''| / c$ 和 $k_{n\rightarrow}'' = \omega''_n |n_{n\rightarrow}''| / c$ ，得到

$$k_{n\leftarrow}'' = k_{n\rightarrow}'' \quad (C198)$$

接着由于双负材料中 $\pi/2 < \alpha''_{n\leftarrow} = \langle \tilde{\mathbf{S}}''_{n\leftarrow}, \mathbf{k}_{n\leftarrow}'' \rangle < 3\pi/2$ ， $\tilde{\mathbf{S}}''_{n\leftarrow}$ 与 $\mathbf{k}_{n\leftarrow}''$ 近似反向，因此若 $\tilde{\mathbf{S}}''_{n\leftarrow}$ 指向第三象限，则 $\mathbf{k}_{n\leftarrow}''$ 将指向第一象限，与反射光的波法线方向夹角小于 $\pi/2$ ，但由于 $\theta' \neq \theta''$ 而不同向；这也可通过 $\gamma''_{n\rightarrow} = \langle \mathbf{k}'_{n\rightarrow T}, \mathbf{n} \rangle > \pi/2$ ，所导致的 $\gamma''_{n\leftarrow} = \langle \mathbf{k}_{n\leftarrow}'' \mathbf{n} \rangle = \pi - \gamma''_{n\rightarrow} < \pi/2$ 看出；因此 $\mathbf{k}_{n\leftarrow T}''$ 与 $\mathbf{k}'_{n\leftarrow T}$ 、 $\mathbf{k}'_{n\rightarrow T}$ 、 $\mathbf{k}_{n\rightarrow T}''$ 均同向，且有 $\mathbf{k}_{n\leftarrow N}''$ 与 $\mathbf{k}'_{n\leftarrow N}$ 、 $\mathbf{k}'_{n\rightarrow N}$ 、 \mathbf{n} 同向，与 $\mathbf{k}_{n\rightarrow N}''$ 反向，即：

$$\widehat{\mathbf{k}_{n\leftarrow T}''} = \widehat{\mathbf{k}_{n\rightarrow T}''} \quad , \quad \widehat{\mathbf{k}_{n\leftarrow N}''} = -\widehat{\mathbf{k}_{n\rightarrow N}''} = \mathbf{n} \quad (C199)$$

结合以上左右手材料中折射光各分量的相对方向，利用 $\gamma''_{n\leftarrow} = \pi - \gamma''_{n\rightarrow}$ ， \sin 作用于两边并乘以 $k_{n\leftarrow}'' = k_{n\rightarrow}''$ 和 $\widehat{\mathbf{k}_{n\leftarrow T}''} = \widehat{\mathbf{k}_{n\rightarrow T}''}$ 得 $\mathbf{k}_{n\leftarrow T}'' = k_{n\leftarrow}'' \sin \gamma''_{n\leftarrow} \widehat{\mathbf{k}_{n\leftarrow T}''} = k_{n\rightarrow}'' \sin \gamma''_{n\rightarrow} \widehat{\mathbf{k}_{n\rightarrow T}''} = \mathbf{k}_{n\rightarrow T}''$ ； \cos 作用于两边并乘以 $k_{n\rightarrow}'' = k_{n\leftarrow}''$ 和 \mathbf{n} 后，得 $\mathbf{k}_{n\leftarrow N}'' = k_{n\leftarrow}'' \cos \gamma''_{n\leftarrow} \mathbf{n} = k_{n\rightarrow}'' \cos(\pi - \gamma''_{n\rightarrow}) \mathbf{n} = -k_{n\rightarrow}'' \cos \gamma''_{n\rightarrow} \mathbf{n} = -\mathbf{k}_{n\rightarrow N}''$ 。总地便有与式(C189)相同的：

$$\mathbf{k}_{n\leftarrow T}'' = \mathbf{k}_{n\rightarrow T}'' \quad , \quad \mathbf{k}_{n\leftarrow N}'' = -\mathbf{k}_{n\rightarrow N}'' \quad (C200)$$

可见式(C200)与式(C184)截然相反，即双负与双正介质中波矢分量的关系与能流分量的关系相反，以至于二者的波矢方向是关于界面对称，而能流方向关于法线对称：

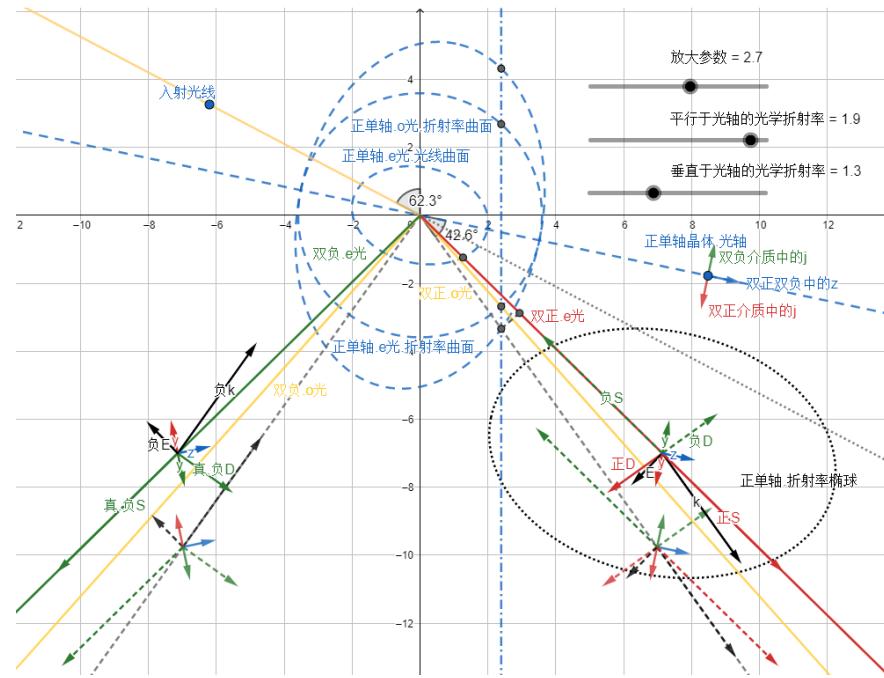


图 C1 双正和双负正单轴晶体中的入射光、正常和反常折射时的 o 光 e 光

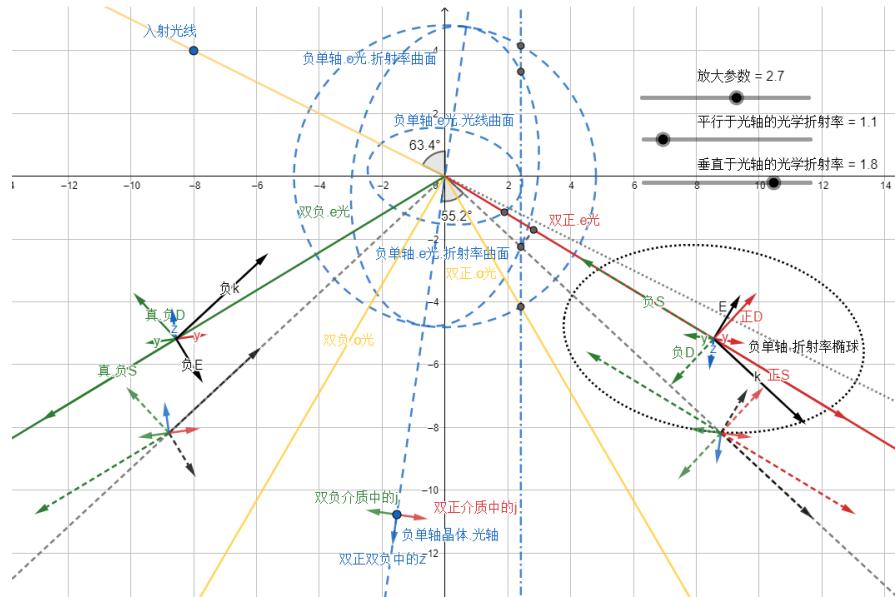


图 C2 双正和双负负单轴晶体中的入射光、正常和反常折射时的 o 光 e 光

⑤. 双正和双负材料中的折射与双折射、反射与双反射，本质上均是电磁场的边值关系、及其所导致的能流密度关系，Snell 定律和波矢的切向关系只是其反映之一，不是其本质

可见此时发生的折射或反常折射，其向下传播的折射光完全是能流密度在参与，

而此时反常介质中的波法线方向与界面外法线方向夹角为锐角，是从负折射率材料中向界面和界面外的方向传播的，是斜向上传播的；而正折射率介质中的折射光的能流密度矢量和波矢均朝远离界面深入介质的方向斜向下传播的。

当然这也进一步佐证了材料在此一般情形下并未向外辐射能量，而不违背能量守恒定律，反而是相速度和波阵面在朝着晶面靠近，尝试着传播到晶面外去，但可能在界面上终止了，或发生了第二次与法线同侧的出射界面的非正常折射和介质内的正常反射。如果二次折射和反射发生了，则所得的材料内的二次反射光的波法线与正折射率介质中的一次折射光同向，光线方向与负折射率介质中一次折射光一同构成介质内的光线反射；而二次折射光的波法线与一次入射光同向，能流密度与之反向。而所可能发生的第三次折射和反射与第一次折射反射相同，因此可纳入第一次折反射过程，构成一个闭环，也就没有第四次（可纳入第二次）以及之后的了。

负折射率材料是否有上述多次折反射，取决于参与折射和反射的到底是波矢还是能流密度：若是能流密度，则上述多次折射反射不成立，反之成立。从 Snell 定律来看，似乎是波矢参与的折射和反射，但从电磁场边界条件来看，又是能流密度矢量参与的折射和反射。而从右手材料虽只有一条折射光的实验现象中又无法证明这一点，只有当左手材料折射光也只有与法线同侧的一条时，就能从实验上证明是能流实实在在参与的折射，以至波矢关系式只是辅助满足相位匹配和时变电磁场边值关系而已。

⑥. 由于实际参与折射的是波射线而不是波法线，因此即使对于右手材料和正常的单折射，本质上也需要先考察其能流方向，再确定其波矢方向，进而确定能流方向

以上考虑启发我们即使对于右手材料，也得先确定其 \tilde{S}_n'' 再确定其 k_n'' ；否则仅凭 Snell 定律是无法确定 k_n'' 的，因为 Snell 定律只能确定四条 k_n'' 所可能存在于射线方向。所以还需要用上缺失的电磁场定矢的边值关系信息，以先确定 \tilde{S}_n'' 的绝对位置，才能得到完备的解。对于右手材料中 \tilde{S}_n'' 的位置的确定，右手材料下表面的剩余四项待确定的参数为 $E_{0nN}''|_{S^-}, H_{0nN}''|_{S^-}, D_{0nT}''|_{S^-}, B_{0nT}''|_{S^-}$ ，虽然由于界面两侧折射率或介电常数张量不同，导致它与介质外的入射光的这四个相应量之间的关系没有与左手介质中的折射光的相应量那么明确，但由于界面两侧的折射率均为正，所以可导出右手材料中折射光的这四项参数与入射光的相应参数近似同向、点乘为正、夹角小于九十度，这一点也可通过菲涅尔 Fresnel 公式看出。因此右手介质中的折射光光线 \tilde{S}_n'' 与其入射光光线指向同一象限、位于法线异侧。由于离散角 $\langle \tilde{S}_n'', k_n'' \rangle$ 小于九十度的缘故，Snell 定律所确定的四条 k_n'' 所指向的四个象限中与 \tilde{S}_n'' 所指向的相同那个象限中的 k_n'' 便是

真实的 \mathbf{k}_n'' ，于是此时在通过菲涅尔公式确定了 $\tilde{\mathbf{S}}_n''$ 的大致方向后，便可根据离散角 $\langle \tilde{\mathbf{S}}_n'', \mathbf{k}_n'' \rangle$ 确定 \mathbf{k}_n'' 的大致方向，以至于通过 Snell 定律确定 \mathbf{k}_n'' 的精确方向，并再次通过离散角 $\langle \tilde{\mathbf{S}}_n'', \mathbf{k}_n'' \rangle$ 公式确定右手材料中 $\tilde{\mathbf{S}}_n''$ 的精确方向。当然另一种方法也可以通过菲涅尔公式配合 Snell 公式、折射率曲面、光线曲面三者来共同确定（见图 C1、C2），但仍需要前者给出能流的大致朝向，再后三者给出能流的精确方向。

从这里也可看出，实际参与折射的，是能流密度矢量，而不是波矢或相速度，能流作为能量、信息和场即物质的载体，是客观存在的实体，而波矢和相速度只是等相位面罢了，是描述该物质的属性之一，是为描述其属性而存在的标签，而不是实际参与相互作用的客体、客观对象——比如菲涅尔公式及其所导出的关于反射波的入射角之布儒斯特角、光疏到光密介质的反射波的半波损失等，无不彰显了波矢之外的场的实际存在及其重要性，所以 Snell 定律相对于场和能流相关的公式，最多只算个约束条件或相位匹配条件。同时这也启发我们，折射率是否可为负不是必要的，因为即使当介电常数为负，折射率也可不定义为负，此时斯涅耳定律仍成立，且双负材料中的能流密度矢量和波矢量仍然可以被正确地描述，且可被更好、更准确地描述。

附C1.2.10 电磁均各向同性的介质中的寻常电磁波

在电磁均各向同性的线性介质中，还额外有

$$\begin{aligned} \mathbf{D}_n(\mathbf{r}) &= \overset{\rightarrow}{\epsilon}_n^{(1)} \cdot \mathbf{E}_n(\mathbf{r}) = \epsilon_n^{(1)} \mathbf{E}_n(\mathbf{r}) \\ \nabla \cdot \mathbf{E}_n(\mathbf{r}) &= i \mathbf{k}_n \cdot \mathbf{E}_n(\mathbf{r}) = 0 \\ \mathbf{H}_{n\perp}(\mathbf{r}) &= -v_n \epsilon_n^{(1)} \mathbf{E}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n = v_n^2 \epsilon_n^{(1)} \mathbf{B}_n(\mathbf{r}) = \frac{1}{\mu_0 \mu_{rn}^{(1)}} \mathbf{B}_n(\mathbf{r}) = \frac{1}{\mu_0 n_{nm} |n_{nm}|} \mathbf{B}_n(\mathbf{r}) \end{aligned} \quad (\text{C201})$$

则此时有：

$$\left\{ \begin{array}{l} \mathbf{E}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n = 0 \\ \mathbf{H}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n = 0 \\ \mathbf{E}_{n\perp}(\mathbf{r}) = \frac{1}{\epsilon_0 n_{ne} |n_{ne}|} \mathbf{D}_n(\mathbf{r}) = \frac{\epsilon_n^{(1)}}{\epsilon_0 \epsilon_{rn}^{(1)}} \mathbf{E}_n(\mathbf{r}) = \mathbf{E}_n(\mathbf{r}) \\ \mathbf{H}_{n\perp}(\mathbf{r}) = \frac{1}{\mu_0 n_{nm} |n_{nm}|} \mathbf{B}_n(\mathbf{r}) = \frac{\mu_n^{(1)}}{\mu_0 \mu_{rn}^{(1)}} \mathbf{H}_n(\mathbf{r}) = \mathbf{H}_n(\mathbf{r}) \end{array} \right. \quad (\text{C202})$$

其中一式与三式、二式与四式互相印证。但却因此少了额外的两组叉乘信息。现 (C93) 左侧式变为：

$$\left\{ \begin{array}{l} \mathbf{E}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n = 0 \\ \mathbf{H}_n(\mathbf{r}) \cdot \hat{\mathbf{k}}_n = 0 \\ \mathbf{E}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n = -v_n \mathbf{B}_n(\mathbf{r}) = -v_n \mu_n^{(1)} \mathbf{H}_n(\mathbf{r}) = -\sqrt{\frac{\mu_n^{(1)}}{\epsilon_n^{(1)}}} \mathbf{H}_n(\mathbf{r}) \\ \mathbf{H}_n(\mathbf{r}) \times \hat{\mathbf{k}}_n = v_n \mathbf{D}_n(\mathbf{r}) = v_n \epsilon_n^{(1)} \mathbf{E}_n(\mathbf{r}) = \sqrt{\frac{\epsilon_n^{(1)}}{\mu_n^{(1)}}} \mathbf{E}_n(\mathbf{r}) \end{array} \right. \quad (C203)$$

该式中三式与四式表达了同一个意义，因此只需存在其中一个；且此时(C188)的三式与(C94)最右侧的四式相互印证；且此时

$$E_{0n} = |\mathbf{E}_n(\mathbf{r})| = \sqrt{\frac{\mu_n^{(1)}}{\epsilon_n^{(1)}}} |\mathbf{H}_n(\mathbf{r})| = H_{0n} \quad (C204)$$

此时 $\tilde{\mathbf{S}}_{n\perp} = -(\tilde{\mathcal{E}}_n \cdot \mathbf{v}_n) \tilde{\mathcal{D}}_n = \mathbf{0}$ ， $\tilde{\mathbf{S}}_n = \tilde{\mathbf{S}}_{n\parallel} = (\tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n) \mathbf{v}_n$ ， 对比 $\tilde{\mathbf{S}}_n = |\tilde{w}_n| \mathbf{v}_{nr}$ 得此时 $\tilde{w}_n = \tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n = \epsilon_n^{(1)} \tilde{\mathcal{E}}_n^2 > 0$ 、 $\mathbf{v}_{nr} = \mathbf{v}_n$ 、 $\langle \tilde{\mathbf{S}}_n, \hat{\mathbf{k}}_n \rangle = \langle \mathbf{v}_{nr}, \mathbf{v}_n \rangle = \alpha_n = 0$ 。 并有

$$\mathbf{I}_n = \frac{\mathcal{S}_n}{2} = \frac{1}{2} \mathcal{E}_n \times \mathcal{H}_n = 2 \mathbf{E}_{0n} \times \mathbf{H}_{0n} = 2 \sqrt{\frac{\epsilon_n^{(1)}}{\mu_n^{(1)}}} E_{0n}^2 \hat{\mathbf{k}}_n \quad (C205)$$

至此，线性和非线性、磁各向同性、电各向同性、电各向异性晶体中，波动方程的解的情况基本明朗了。

附录 D 线性介质的二阶极化率张量为对称张量

据坡印廷定理，无损介质中，有以下能量守恒过程：单位时间内流入系统的能量，等于单位时间内场能的增加量，与单位时间内带电体动能增量：

$$\frac{d\tilde{W}_i}{dt} = \frac{d\tilde{W}}{dt} + \frac{d\tilde{E}_k}{dt} \quad (D1)$$

其中：

$$\frac{d\tilde{W}_i}{dt} = - \oint_S \tilde{\mathbf{S}} \cdot d\sigma \quad (D2)$$

$$\begin{aligned} \frac{d\tilde{E}_k}{dt} - \frac{d\tilde{F} - dV}{dL - dt} \rightarrow \frac{d\tilde{F} \cdot dL}{dt} &= \tilde{dF} \cdot \tilde{v} = [\tilde{dF}_c + \tilde{dF}_m] \cdot \tilde{v} = [(\tilde{\rho}_f dV) \tilde{\mathcal{E}} + (\tilde{\mathcal{J}}_f dV) \times \tilde{\mathbf{B}}] \cdot \tilde{v} \\ &= (\tilde{\rho}_f dV) [\tilde{\mathcal{E}} + \tilde{v} \times \tilde{\mathbf{B}}] \cdot \tilde{v} \xrightarrow[\tilde{\mathcal{E}} \cdot \tilde{\rho}_f \tilde{v} dV]{(\tilde{\rho}_f dV) \tilde{\mathcal{E}} \cdot \tilde{v}} \left\{ \begin{array}{l} \frac{d\tilde{F}_c \cdot \tilde{v}}{dt} \\ \tilde{\mathcal{E}} \cdot \tilde{\mathcal{J}}_f dV \end{array} \right. \end{aligned} \quad (D3)$$

即有：

$$-\iiint_V \nabla \cdot \tilde{\mathbf{S}} dV = \frac{d}{dt} \iiint_V (\tilde{w} + \tilde{w}_k) dV = \iiint_V \left(\frac{\partial \tilde{w}}{\partial t} + \tilde{dF}_c \cdot \tilde{v} \right) dV \quad (D4)$$

得坡印廷定理的微分形式：

$$-\nabla \cdot \tilde{\mathbf{S}} = \frac{\partial \tilde{w}}{\partial t} + \tilde{dF}_c \cdot \tilde{v} = \frac{\partial \tilde{w}}{\partial t} + \tilde{\mathcal{E}} \cdot \tilde{\mathcal{J}}_f \quad (D5)$$

利用之，并配合微分形式麦氏方程组的四号方程有：

$$\begin{aligned} -\tilde{\mathcal{H}} \cdot \frac{\partial \tilde{\mathbf{B}}}{\partial t} - \nabla \cdot (\tilde{\mathcal{E}} \times \tilde{\mathcal{H}}) &= \tilde{\mathcal{H}} \cdot (\nabla \times \tilde{\mathcal{E}}) - \nabla \cdot (\tilde{\mathcal{E}} \times \tilde{\mathcal{H}}) \xleftarrow{\nabla \cdot (a \times b) = b \cdot (\nabla \times a) - a \cdot (\nabla \times b)} \tilde{\mathcal{E}} \cdot (\nabla \times \tilde{\mathcal{H}}) \\ \xleftarrow[\nabla \times \tilde{\mathcal{H}} = \tilde{\mathcal{J}}_f + \frac{\partial \tilde{\mathcal{D}}}{\partial t}]{} &\tilde{\mathcal{E}} \cdot (\tilde{\mathcal{J}}_f + \frac{\partial \tilde{\mathcal{D}}}{\partial t}) = -\frac{\partial \tilde{w}}{\partial t} + \tilde{\mathcal{E}} \cdot \frac{\partial \tilde{\mathcal{D}}}{\partial t} - \nabla \cdot \tilde{\mathbf{S}} \end{aligned} \quad (D6)$$

对比得：

$$\tilde{\mathbf{S}} = \tilde{\mathcal{E}} \times \tilde{\mathcal{H}} \quad (D7)$$

$$\frac{\partial \tilde{w}}{\partial t} = \tilde{\mathcal{E}} \cdot \frac{\partial \tilde{\mathcal{D}}}{\partial t} + \tilde{\mathcal{H}} \cdot \frac{\partial \tilde{\mathbf{B}}}{\partial t} \quad (D8)$$

$$d\tilde{w} = \tilde{\mathcal{E}} \cdot d\tilde{\mathcal{D}} + \tilde{\mathcal{H}} \cdot d\tilde{\mathbf{B}} = d\tilde{w}_e + d\tilde{w}_m \quad (D9)$$

因麦氏方程组和坡印廷定理均是普适的，则由二者推出的这三个关系式也是普适的，适用于复合场（非单色波亦可）以及非线性介质。在线性介质中， $\tilde{\mathcal{D}} = \tilde{\mathcal{D}}^{(1)} = \overset{\leftrightarrow}{\epsilon}^{(1)} \cdot \tilde{\mathcal{E}}$ 、 $\tilde{\mathbf{B}} = \tilde{\mathbf{B}}^{(1)} = \overset{\leftrightarrow}{\mu}^{(1)} \cdot \tilde{\mathcal{H}}$ ，此时可对

$$d\tilde{w} \xrightarrow[d\overset{\leftrightarrow}{\mu}^{(1)} = \overset{\leftrightarrow}{\mathbf{0}}]{d\overset{\leftrightarrow}{\epsilon}^{(1)} = \overset{\leftrightarrow}{\mathbf{0}}} \tilde{\mathcal{E}} \cdot \overset{\leftrightarrow}{\epsilon}^{(1)} \cdot d\tilde{\mathcal{E}} + \tilde{\mathcal{H}} \cdot \overset{\leftrightarrow}{\mu}^{(1)} \cdot d\tilde{\mathcal{H}}$$

$$\frac{\text{设 } \overset{\rightarrow}{\boldsymbol{\epsilon}}^{(1)}, \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)} \text{ 是对称张量, 则满足交换律}}{\tilde{\boldsymbol{\epsilon}} \cdot \overset{\rightarrow}{\boldsymbol{\epsilon}}^{(1)} = \overset{\rightarrow}{\boldsymbol{\epsilon}}^{(1)} \cdot \tilde{\boldsymbol{\epsilon}}, \tilde{\boldsymbol{\mu}} \cdot \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)} = \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)} \cdot \tilde{\boldsymbol{\mu}}} \overset{\rightarrow}{\boldsymbol{\epsilon}}^{(1)} \cdot \tilde{\boldsymbol{\epsilon}} + \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)} \cdot \tilde{\boldsymbol{\mu}} \quad (\text{D10})$$

于任意场点处积分得:

$$\tilde{w} = \frac{1}{2} \tilde{\boldsymbol{\mathcal{D}}} \cdot \tilde{\boldsymbol{\mathcal{E}}} + \frac{1}{2} \tilde{\boldsymbol{\mathcal{B}}} \cdot \tilde{\boldsymbol{\mathcal{H}}} \quad (\text{D11})$$

此时

$$\tilde{w}_e = \frac{1}{2} \tilde{\boldsymbol{\mathcal{E}}} \cdot \tilde{\boldsymbol{\mathcal{D}}} \quad (\text{D12})$$

$$\tilde{w}_m = \frac{1}{2} \tilde{\boldsymbol{\mathcal{H}}} \cdot \tilde{\boldsymbol{\mathcal{B}}} \quad (\text{D13})$$

于是:

$$d\tilde{w}_e = d\left(\frac{1}{2} \tilde{\boldsymbol{\mathcal{E}}} \cdot \tilde{\boldsymbol{\mathcal{D}}}\right) = \frac{1}{2} \tilde{\boldsymbol{\mathcal{E}}} \cdot d\tilde{\boldsymbol{\mathcal{D}}} + \frac{1}{2} \tilde{\boldsymbol{\mathcal{D}}} \cdot d\tilde{\boldsymbol{\mathcal{E}}} \quad (\text{D14})$$

对比普适的 $d\tilde{w}_e = \tilde{\boldsymbol{\mathcal{E}}} \cdot d\tilde{\boldsymbol{\mathcal{D}}}$, 得:

$$\tilde{\boldsymbol{\mathcal{D}}} \cdot d\tilde{\boldsymbol{\mathcal{E}}} = \sum_{u_1}^{\{i,j,k\}} \tilde{\boldsymbol{\mathcal{D}}}_{u_1} d\tilde{\boldsymbol{\mathcal{E}}}_{u_1} = \sum_{u_1 u_2}^{\{i,j,k\}} \varepsilon_{u_1 u_2}^{(1)} \tilde{\boldsymbol{\mathcal{E}}}_{u_2} \cdot d\tilde{\boldsymbol{\mathcal{E}}}_{u_1} = \sum_{u_1 u_2}^{\{i,j,k\}} \tilde{\boldsymbol{\mathcal{E}}}_{u_1} \cdot \varepsilon_{u_1 u_2}^{(1)} d\tilde{\boldsymbol{\mathcal{E}}}_{u_2} = \tilde{\boldsymbol{\mathcal{E}}} \cdot d\tilde{\boldsymbol{\mathcal{D}}} \quad (\text{D15})$$

即有 $\varepsilon_{u_1 u_2}^{(1)} = \varepsilon_{u_2 u_1}^{(1)}$, 即印证了之前的假设: $\overset{\rightarrow}{\boldsymbol{\epsilon}}^{(1)}, \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)}$ 是对称张量, 且其与任何三维矢量相乘均满足交换律; 则此时 $\overset{\rightarrow}{\boldsymbol{\epsilon}}_r^{(1)}, \overset{\rightarrow}{\boldsymbol{\mu}}_r^{(1)}, \overset{\rightarrow}{\boldsymbol{\chi}}_e^{(1)}, \overset{\rightarrow}{\boldsymbol{\chi}}_m^{(1)}$ 均为对称张量。

但在附 E 中类似的过程会给出, 式(D15)中对复色场不能先将 $\tilde{\boldsymbol{\mathcal{E}}}, \tilde{\boldsymbol{\mathcal{D}}}$ 的关系代入, 第二步再将复色场 $\tilde{\boldsymbol{\mathcal{E}}}, \tilde{\boldsymbol{\mathcal{D}}}$ 按实数形式单色场级数展开, 或进一步将其分解为正负频复数形式单色子波, 即将复色场按复数形式的单色场级数展开。因为第一步中的 $\varepsilon_{u_1 u_2}^{(1)}$ 因是非色散的而没有实际意义, 所以正确的做法是先将复色场 $\tilde{\boldsymbol{\mathcal{E}}}, \tilde{\boldsymbol{\mathcal{D}}}$ 级数展开, 再代入

$$\tilde{\boldsymbol{\mathcal{D}}}_{nu_1} = \sum_{u_2}^{\{i,j,k\}} \varepsilon_{nu_1 u_2}^{(1)} \tilde{\boldsymbol{\mathcal{E}}}_{nu_2} \quad \text{或} \quad \tilde{\boldsymbol{\mathcal{D}}}_{nu_1} = \sum_{u_2}^{\{i,j,k\}} \varepsilon_{nu_1 u_2}^{(1)} \tilde{\boldsymbol{\mathcal{E}}}_{nu_2} \quad (\text{D16})$$

当然, 由于这里不需要也不能给出“复色非线性系数”张量的转置对称性, 所以可直接将式(D15)所有量单色化后作为结论, 即单色化后的 $\overset{\rightarrow}{\boldsymbol{\epsilon}}^{(1)}(\omega_n), \overset{\rightarrow}{\boldsymbol{\mu}}^{(1)}(\omega_n), \overset{\rightarrow}{\boldsymbol{\epsilon}}_r^{(1)}(\omega_n), \overset{\rightarrow}{\boldsymbol{\mu}}_r^{(1)}(\omega_n), \overset{\rightarrow}{\boldsymbol{\chi}}_e^{(1)}(\omega_n), \overset{\rightarrow}{\boldsymbol{\chi}}_m^{(1)}(\omega_n)$ 是对称张量。因此即使在色散介质中, 只要介质仍然是无损的, 那么仍有单色化后的一阶非线性系数、甚至高阶非线性系数, 均都是对称张量, 而高阶单色非线性系数的对称是全/任意置换对称外加全取反对称, 将在附录 E 中引入。

注意即使在色散介质中, 讨论复色非线性系数仍没有意义, 即使此时它们看上去是单色非线性系数的线性叠加而可能也是对称的。不过确实有相对于介质固有频率而言甚低频入射光下的近似无色散介质, 此时介质的各阶非线性系数与频率近似无关,

但最好也不称其为各阶复色非线性系数，而只是与频率无关的常数而恰好普适于所有频率的非线性过程罢。

附录 E 各阶非线性光学系数的性质

为简化接下来及正文计算所需用到的表述，现将式(2.1)与式(2.3)进行简写，由于 $\omega_{m_{\pi(i)}}$ 的写法无法一目了然地看出该频率属于哪束光，则先将其中的子角标写作 $\pi(1) := i_1$ ；而 m 因已完成了区别两套频谱、引入光束的频率简并性质等历史使命而均不需要再额外强调，因此省略所有物理量的角标 m ，并将其旗下各级子角标升级为角标。这种简写总是可行的，因为 $\omega_{m_{\pi(a)}} = \omega_{m_{i_a}} = \omega_{i_a}$ 。于是有：

$$\begin{aligned} \tilde{\mathbf{P}}_n(\mathbf{r}, t) &= \tilde{\mathbf{P}}(\omega_n) = \varepsilon_0 \left[\sum_{\omega_n = \omega_{i_1}} \overrightarrow{\chi}_e^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}(\omega_{i_1}) + \sum_{\omega_n = \omega_{i_1} + \omega_{i_2}} D_{i_1, i_2} \overrightarrow{\chi}_e^{(2)}(\omega_n; \omega_{i_1}, \omega_{i_2}) \right. \\ &\quad : \tilde{\mathbf{E}}(\omega_{i_1}) \tilde{\mathbf{E}}(\omega_{i_2}) + \dots + \sum_{\omega_n = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_l}} D_{i_1, i_2, \dots, i_l} \overrightarrow{\chi}_e^{(l)}(\omega_n; \overbrace{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}}^l) \\ &\quad \left. \vdots \tilde{\mathbf{E}}(\omega_{i_1}) \tilde{\mathbf{E}}(\omega_{i_2}) \cdots \tilde{\mathbf{E}}(\omega_{i_l}) + \dots \right] \end{aligned} \quad (E1)$$

以及：

$$\begin{aligned} \tilde{\mathbf{P}}_{u_1}(\omega_n) &= \varepsilon_0 \left[\sum_{\omega_n = \omega_{i_1}} \sum_{u_2}^{\{i, j, k\}} \chi_{u_1 u_2}^{(1)}(\omega_n) \tilde{\mathbf{E}}_{u_2}(\omega_{i_1}) + \sum_{\omega_n = \omega_{i_1} + \omega_{i_2}} D_{i_1, i_2} \right. \\ &\quad \cdot \sum_{u_2 u_3}^{\{i, j, k\}} \chi_{u_1 u_2 u_3}^{(2)}(\omega_n; \omega_{i_1}, \omega_{i_2}) \tilde{\mathbf{E}}_{u_2}(\omega_{i_1}) \tilde{\mathbf{E}}_{u_3}(\omega_{i_2}) + \dots + \sum_{\omega_n = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_l}} D_{i_1, i_2, \dots, i_l} \\ &\quad \left. \cdot \sum_{u_2 u_3 \cdots u_{l+1}}^{\{i, j, k\}} \chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \overbrace{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}}^l) \tilde{\mathbf{E}}_{u_2}(\omega_{i_1}) \tilde{\mathbf{E}}_{u_3}(\omega_{i_2}) \cdots \tilde{\mathbf{E}}_{u_{l+1}}(\omega_{i_l}) + \dots \right] \end{aligned} \quad (E2)$$

附E1 各阶 $\overrightarrow{\chi}_e^{(l)}$ 的各分量和其复共轭之间，与各单色场遵循相同规则

$$\overrightarrow{\chi}_e^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l})^* = \overrightarrow{\chi}_e^{(l)}(-\omega_n; -\omega_{i_1}, -\omega_{i_2}, \dots, -\omega_{i_l}) \quad (E3)$$

对于非线性光学效应的分量表达式(E2)，也有类似的：

$$\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l})^* = \chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(-\omega_n; -\omega_{i_1}, -\omega_{i_2}, \dots, -\omega_{i_l}) \quad (E4)$$

其中要注意，式(E2)左右两侧的非线性系数下标取值的排列需要一致（如两侧均为 $u_1 u_2 u_3 \cdots u_{l+1} = kij \cdots i$ 或 $213 \cdots 2$ 、 $xzy \cdots z$ 等），且对所有排列均成立。

其证明过程已在正文式(2.17)紧邻的下一段给出。

附E2 简并因子所暗示的非线性系数的内秉的(角标)排列/置换对称性

考虑式(E2)中的第 l 项 (第 l 阶非线性光学效应)

$$\sum_{\omega_n = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_l}} D_{i_1, i_2, \dots, i_l} \sum_{u_2 u_3 \cdots u_{l+1}}^{\{i, j, k\}} \chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)} (\omega_n; \overbrace{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}}^l) \cdot \overbrace{\tilde{E}_{u_2}(\omega_{i_1}) \tilde{E}_{u_3}(\omega_{i_2}) \cdots \tilde{E}_{u_{l+1}}(\omega_{i_l})}^l \quad (E5)$$

将上式中第任意束光如第 l 束光的相应下标如 u_{l+1}, i_l , 分别同时与另某束光如第 2 束光的相应下标 u_3, i_2 调换位置, 得到的式子的值理应与原式的值相等, 因为光线或光线标签交换顺序的同时, 可以让其下标 i_l 与 i_2 的取值 ($i_l \neq i_2$) 同时交换为对方的值, 那么 i_l 与 i_2 的值在空间位置上是没有交换的, 进而不会影响公式的结果。这也是简并因子 D_{i_1, i_2, \dots, i_l} 的物理意义所在和出现的原因: 当公式对和频光的非线性过程的多种描述均等价时, 这些多种描述都该算进同一个和频光的非线性过程中去。简而言之, 当公式中引入了简并因子 D_{i_1, i_2, \dots, i_l} 后, 描述同一个非线性过程的公式便不再关心其排列地对组合相同的混频光场给出相同结果, 所以有:

$$\sum_{\omega_n = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_2}} D_{i_1, i_2, \dots, i_2} \sum_{u_2 u_{l+1} \cdots u_3}^{\{i, j, k\}} \chi_{u_1 u_2 u_{l+1} \cdots u_3}^{(l)} (\omega_n; \overbrace{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_2}}^l) \cdot \overbrace{\tilde{E}_{u_2}(\omega_{i_1}) \tilde{E}_{u_{l+1}}(\omega_{i_2}) \cdots \tilde{E}_{u_3}(\omega_{i_2})}^l \quad (E6)$$

其中, 因加法交换律有 $\omega_n = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_2} = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_l}$, 所以

$$\sum_{\omega_n = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_2}} \Leftrightarrow \sum_{\omega_n = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_l}} \quad (E7)$$

又因正文式(2.2)而有:

$$D_{i_1, i_2, \dots, i_2} = \frac{l!}{\sum_m \text{card}\{\overline{m_i | m_i = m}\}!} = D_{i_1, i_2, \dots, i_l} \quad (E8)$$

因下面二者均为取遍 $u_2 u_3 \cdots u_{l+1}$ 的所有排列, 所以有

$$\sum_{u_2 u_{l+1} \cdots u_3}^{\{i, j, k\}} \Leftrightarrow \sum_{u_2 u_3 \cdots u_{l+1}}^{\{i, j, k\}} \quad (E9)$$

因乘法交换律, 有

$$\tilde{E}_{u_2}(\omega_{i_1}) \tilde{E}_{u_3}(\omega_{i_2}) \cdots \tilde{E}_{u_{l+1}}(\omega_{i_l}) = \tilde{E}_{u_2}(\omega_{i_1}) \tilde{E}_{u_{l+1}}(\omega_{i_l}) \cdots \tilde{E}_{u_3}(\omega_{i_2}) \quad (E10)$$

将上述(E7-10)式全代入式(E6), 得到

$$\sum_{\omega_n = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_l}} D_{i_1, i_2, \dots, i_l} \sum_{u_2 u_3 \cdots u_{l+1}}^{\{i, j, k\}} \chi_{u_1 u_2 u_{l+1} \cdots u_3}^{(l)} (\omega_n; \overbrace{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}}^l) \overbrace{\tilde{E}_{u_2}(\omega_{i_1}) \tilde{E}_{u_3}(\omega_{i_2}) \cdots \tilde{E}_{u_{l+1}}(\omega_{i_l})}^l \quad (E11)$$

对比式(E11)与式(E5)，便得到

$$\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}) = \chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}) \quad (\text{E12})$$

不失普遍地，该式对于调换任意一对 $1 \leq a, b \leq l$ 的 u_{a+1}, i_a 与 u_{b+1}, i_b 均成立（其中的 i_a, i_b 也可描述为 $\omega_{i_a}, \omega_{i_b}$ ）。

由于该置换对称性是因描述非线性过程的公式在创设之初加入了简并因子，相同组合但不同排列的光线简并为同一光线序列的思想，即简并因子存在之本身，便将同时有此结论。这是纯粹数学上的结论和假设，不涉及真实的物理过程，所以称之为非线性光学系数的内秉性质。

附E3 无损介质中的(角标)全(/任意)置换(/对换)对称性

附E3.1 无损介质中的(频率)全取反对称性

附录 C 通过给出复二阶介电常数张量下的波动方程会解出倏逝波，即相应线性介质为有能量损耗的介质，证明了复一阶(非)线性光学系数将导致介质有损，取其逆否命题可知，无损线性介质中的二阶极化率张量的各分量为实数。

站在线性介质中的结论的基础上，对于非线性介质，我们认为该结论可以外推，即无损非线性介质中的各阶非线性光学系数的各分量均为实数，这可以通过考察和对比复、实各阶非线性系数所对应的波动方程的解的不同之处，以及线性和非线性介质中复非线性系数所导致的波动方程的解的共同之处，来得到。因此无损介质中各阶非线性光学系数的复共轭等于其本身，即有：

$$\begin{aligned} \overset{\rightarrow \rightarrow \dots \rightarrow}{\chi_e^{(l)}}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}) &= \overset{\rightarrow \rightarrow \dots \rightarrow}{\chi_e^{(l)}}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}) \\ \chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l})^* &= \chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}) \end{aligned} \quad (\text{E13})$$

将其对比式(E3-4)，无损介质中则还有原料子波与合成极化波全正频与全负频^①所对应的非线性系数相同：

$$\begin{aligned} \overset{\rightarrow \rightarrow \dots \rightarrow}{\chi_e^{(l)}}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}) &= \overset{\rightarrow \rightarrow \dots \rightarrow}{\chi_e^{(l)}}(-\omega_n; -\omega_{i_1}, -\omega_{i_2}, \dots, -\omega_{i_l}) \\ \chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}) &= \chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(-\omega_n; -\omega_{i_1}, -\omega_{i_2}, \dots, -\omega_{i_l}) \end{aligned} \quad (\text{E14})$$

^① 注意这里以及其余大多数地方的“全负”均是指全取相反数，但各个频率取值是正是负都可。

要注意(E3-4)、(E13-14)三者的区别，且后两者只在无损介质中成立。

附E3.2 无损介质中的跨等号置换对称性

现根据 Boyd_R.W.在其所著的《Nonlinear Optics》的第 1.5.4 节，可知电场能的时间平均值，即复色电场的内能为：

$$U = \overline{\tilde{w}_e} = \frac{1}{2} \overline{\tilde{\mathcal{E}} \cdot \tilde{\mathcal{D}}} \quad (\text{E15})$$

在线性介质中，类似附录 D 的单色化后的(D15)地有：

$$U_n = \overline{\tilde{w}_{ne}} = \frac{1}{2} \overline{\tilde{\mathcal{E}}_n \cdot \tilde{\mathcal{D}}_n} = \frac{1}{2} \overline{\sum_{u_1}^{\{i,j,k\}} \tilde{\mathcal{E}}_{nu_1} \tilde{\mathcal{D}}_{nu_1}} = \frac{1}{2} \overline{\sum_{u_1 u_2}^{\{i,j,k\}} \tilde{\mathcal{E}}_{nu_1} \cdot \mathcal{E}_{nu_1 u_2}^{(1)} \tilde{\mathcal{E}}_{nu_2}} \quad (\text{E16})$$

现据 $\tilde{\mathcal{E}}_n = \tilde{E}_n + \tilde{E}_n^*$ 标量(分量)化后的 $\tilde{\mathcal{E}}_{nu_1} = \tilde{E}_{nu_1} + \tilde{E}_{nu_1}^*$ ， $u_1 = i, j, k$ ，有：

$$U_n = \frac{1}{2} \overline{\sum_{u_1 u_2}^{\{i,j,k\}} (\tilde{E}_{nu_1} + \tilde{E}_{nu_1}^*) \cdot \mathcal{E}_{nu_1 u_2}^{(1)} (\tilde{E}_{nu_2} + \tilde{E}_{nu_2}^*)} = \frac{1}{2} \overline{\sum_{u_1 u_2}^{\{i,j,k\}} \mathcal{E}_{nu_1 u_2}^{(1)} (\tilde{E}_{nu_1} \tilde{E}_{nu_2}^* + \tilde{E}_{nu_1}^* \tilde{E}_{nu_2})} \quad (\text{E17})$$

其中 $\overline{\tilde{E}_{nu_1} \tilde{E}_{nu_2}} = 0$ 、 $\overline{\tilde{E}_{nu_1}^* \tilde{E}_{nu_2}^*} = 0$ 的理由，可参见(C72-76)，简单来说即每一对的频率之和不为零。同样因 $\tilde{E}_{nu_1}, \tilde{E}_{nu_2}^*$ 、 $\tilde{E}_{nu_1}^*, \tilde{E}_{nu_2}$ 两两频率相反，而有 $\overline{\tilde{E}_{nu_1} \tilde{E}_{nu_2}^*} = E_{nu_1} E_{nu_2}^*$ 、 $\overline{\tilde{E}_{nu_1}^* \tilde{E}_{nu_2}} = E_{nu_1}^* E_{nu_2}$ 。

$$U_n = \frac{1}{2} \overline{\sum_{u_1 u_2}^{\{i,j,k\}} \mathcal{E}_{nu_1 u_2}^{(1)} (E_{nu_1} E_{nu_2}^* + E_{nu_1}^* E_{nu_2})} \quad (\text{E18})$$

接着再因 $\overrightarrow{\mathbf{e}}^{(1)}(\omega_n) = \varepsilon_0 \overrightarrow{\mathbf{e}}_r^{(1)}(\omega_n) = \varepsilon_0 [\overrightarrow{\mathbf{I}} + \overrightarrow{\chi}_e^{(1)}(\omega_n)]$ ，而有 $\varepsilon_{nu_1 u_2}^{(1)} = \varepsilon_0 [\delta_{u_1 u_2} + \chi_{nu_1 u_2}^{(1)}]$ ，于是有：

$$U_n = \frac{1}{2} \varepsilon_0 \overline{\sum_{u_1}^{\{i,j,k\}} (E_{nu_1} E_{nu_1}^* + E_{nu_1}^* E_{nu_1})} + \frac{1}{2} \varepsilon_0 \overline{\sum_{u_1 u_2}^{\{i,j,k\}} \chi_{nu_1 u_2}^{(1)} (E_{nu_1} E_{nu_2}^* + E_{nu_1}^* E_{nu_2})}$$

$$= \varepsilon_0 \overline{\sum_{u_1}^{\{i,j,k\}} E_{nu_1} E_{nu_1}^*} + \varepsilon_0 \overline{\sum_{u_1 u_2}^{\{i,j,k\}} \chi_{nu_1 u_2}^{(1)} E_{nu_1} E_{nu_2}^*} \quad (\text{E19})$$

而根据式(C80)，复色场的能流密度时间平均值满足 $\overline{\tilde{\mathcal{S}}} = \sum_{n=0}^{\infty} \overline{\tilde{\mathcal{S}}_n}$ ，同理根据与式(C77-80)类似的过程，场能密度平均值、电场和磁场能量密度的平均值也均满足：

$$\overline{\tilde{w}} = \sum_{n=0}^{\infty} \overline{\tilde{w}_n} \quad , \quad \overline{\tilde{w}_m} = \sum_{n=0}^{\infty} \overline{\tilde{w}_{nm}} \quad , \quad \overline{\tilde{w}_e} = \sum_{n=0}^{\infty} \overline{\tilde{w}_{ne}} \quad \text{即} \quad U = \sum_{n=0}^{\infty} U_n \quad (\text{E20})$$

那么便直接有：

$$U = \varepsilon_0 \overline{\sum_{u_1}^{\{i,j,k\}} \sum_{n=0}^{\infty} E_{nu_1} E_{nu_1}^*} + \varepsilon_0 \overline{\sum_{u_1 u_2}^{\{i,j,k\}} \sum_{n=0}^{\infty} \chi_{nu_1 u_2}^{(1)} E_{nu_1} E_{nu_2}^*} \quad (\text{E21})$$

该式与 boyd 的式 1.5.14 相照应，但比他少了系数 1/2，这是因为他式中的 $\sum_{n=0}^{\infty}$

取的是 $\sum_{n=-\infty}^{\infty}$ ，所以整体比式(E21)多 2 倍，因此需要除以 2；而 boyd 由于熟悉量子力学而对时间平均值采用了 bracket 记号，并对(E21)末项采用了 $E_{nu_2}^* \chi_{nu_1 u_2}^{(1)} E_{nu_1}$ 的写法；我们其余有关各量均一一对应。当然，若仅仅为了推导无损介质中的全置换对称性，不需要像 boyd 那样给出(E21)，只需要(E19)即可，所以这里只是为了展示与 boyd 相得益彰而给出(E21)的，另一方面也暗示前人走了不少弯路，但确实得到了适用最为广泛的结论。

然而以上只考虑了线性介质，在非线性介质中电位移矢量有高阶项

$$\tilde{D}_n = \sum_{\omega_n = \omega_{i_1}} \overleftrightarrow{\epsilon}^{(1)}(\omega_n) \cdot \tilde{E}(\omega_{i_1}) + \sum_{\omega_n = \omega_{i_1} + \omega_{i_2}} D_{i_1, i_2} \overleftrightarrow{\epsilon}^{(2)}(\omega_n) : \tilde{E}(\omega_{i_1}) \tilde{E}(\omega_{i_2}) + \dots \quad (\text{E22})$$

因此(E16)会变得非常复杂。一般的做法已经不再是把式(E22)的分量/标量表达式代入其中，而是直接对电场能量密度 U 或 U_n 做级数展开。在展开的结果中，各阶非线性系数的一阶导数 $\chi_{nu_1 u_2}^{(1)}$ 将作为级数的各次项系数出现，而由于电场能密度 U 和 U_n 是场能密度，属于能量密度，不是物质性的，因此加和起来凑成 U 和 U_n 的各项中的 $\chi_{nu_1 u_2}^{(1)}$ ，满足跨等号置换对称性，即在满足相同的能量守恒过程 $\omega_n = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_l}$ 下，它们的频率分量 ω_{i_a} 与 ω_{i_b} ($1 \leq a, b \leq l$ ；其中一个可以是 ω_n) 可以两两间任意交换位置^①，只要 $\chi_{u_1 u_2 u_3 \dots u_{l+1}}^{(l)}$ 的相关角标 u_{a+1} 与 u_{b+1} (其中一个可以是 u_1) 也跟着一同交换。在这样的描述下，跨等号置换对称性中包含了内秉对称性，是其升级版；但也可将跨等号置换对称性与内秉对称性之间划清楚河汉界，此时跨等号置换对称性只包含跨等号置换一对频率和相应角标的操作，如 $-\omega_{i_1} = -\omega_n + \omega_{i_2} + \dots + \omega_{i_l}$ 及其所对应的 $\chi_{u_2 u_3 \dots u_{l+1}}^{(l)}$ 。既然的 $\chi_{nu_1 u_2}^{(1)}$ 一阶导 $\chi_{nu_1 u_2}^{(1)}$ 满足跨等号置换对称性，则也认为 $\chi_{nu_1 u_2}^{(1)}$ 也满足跨等号置换对称性，如：

$$\begin{aligned} & \chi_{u_1 u_2 u_3 \dots u_{l+1}}^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}) \xleftarrow{\text{跨等号置换对称性}} \chi_{u_3 u_2 u_1 \dots u_{l+1}}^{(l)}(-\omega_{i_2}; \omega_{i_1}, -\omega_n, \dots, \omega_{i_l}) \\ &= \chi_{u_{l+1} u_2 u_3 \dots u_1}^{(l)}(-\omega_{i_l}; \omega_{i_1}, \omega_{i_2}, \dots, -\omega_n) \xleftarrow{\text{内秉对称性}} \chi_{u_{l+1} u_2 u_1 \dots u_3}^{(l)}(-\omega_{i_l}; \omega_{i_1}, -\omega_n, \dots, \omega_{i_2}) \end{aligned} \quad (\text{E23})$$

^① 均处于频率守恒方程右侧的一对频率互换位置，因加法交换律而不改变二者的符号，且本质上这属于内秉对称性；但若涉及与左侧 ω_n 调换位置时，两调换的频率均需反号，因横跨方程等号左右移项需反号，这才是这里的全置换对称性 A 相对于内秉对称性进步了的地方，或者这才是全置换对称性 A 的意义：描述同样一个能量守恒过程的式子可以不一定要是全加和的形式。注意需保证移项后方程左侧仍只有一个频率，而右侧有 l 个，以保证相应非线性系数仍是 l 阶的。

当然，在对 U 或 U_n 进行展开的时候，第 l 阶非线性系数的一阶导数 $\chi_{mu_2}^{(l)'}'$ ，只是作为其各次项系数地是 U 或 U_n 对场振幅 $E_{mu_2}^*$ 的第 $l+1$ 阶偏导 $U^{(l+1)}$ 、 $U_n^{(l+1)}$ ，不能给出其物理含义，需要对比朗道(Landau)等人给出的式 1.5.16 等才能可赋予各次展开项系数的物理意义为相应第 l 阶非线性系数的一阶导数 $\chi_{mu_2}^{(l)'}'$ 。并且要注意只有在无损介质中内能才是态 $E_{mu_2}^*$ 的函数，否则不含或与 $E_{mu_2}^*$ 无关而没有相应偏导数。

附E3.3 无损介质中的全置换对称性

根据式(E12)以及式(E21)，可得：

$$\begin{aligned}\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}) &= \chi_{u_3 u_2 u_1 \cdots u_{l+1}}^{(l)}(\omega_{i_2}; -\omega_{i_1}, \omega_n, \dots, -\omega_{i_l}) \\ &= \chi_{u_{l+1} u_2 u_3 \cdots u_1}^{(l)}(\omega_{i_l}; -\omega_{i_1}, -\omega_{i_2}, \dots, \omega_n) = \chi_{u_{l+1} u_2 u_1 \cdots u_3}^{(l)}(\omega_{i_l}; -\omega_{i_1}, \omega_n, \dots, -\omega_{i_2})\end{aligned}\quad (\text{E24})$$

意味着可以对方程 $\omega_n = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_l}$ 作任何数学操作，包括左右移单项、右侧加法交换律，以及左右两侧均取负这三种操作，只要能保证其左侧是单项，且 $\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}$ 的相应角标 u_{a+1} 与 u_{b+1} 也一同调换即可。

可见全置换对称性包含了跨等号置换对称性、内秉对称性、全取反对称性。

附E4 克莱曼(Kleinman)对称性

当在晶体中参与混频的入射光集的频率均远远低于介质本身的最低共振频率时，即使介质是色散介质，此时也可将其看做无色散的介质，即此时其各阶非线性系数 $\xrightarrow{\rightarrow \cdots \rightarrow} \chi_e^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l})$ 、 $\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l})$ 与待合成的频率 ω_n 无关的同时，也与参与混频的各单色激发场源的频率 $\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}$ 无关。根据前者即 $\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}$ 与 ω_n 无关，即有诸如：

$$\begin{aligned}\chi_{u_3 u_2 u_1 \cdots u_{l+1}}^{(l)}(\omega_{i_2}) &= \chi_{u_3 u_2 u_1 \cdots u_{l+1}}^{(l)}(\omega_n) \\ \chi_{u_{l+1} u_2 u_3 \cdots u_1}^{(l)}(\omega_{i_l}) &= \chi_{u_{l+1} u_2 u_3 \cdots u_1}^{(l)}(\omega_n) \\ \chi_{u_{l+1} u_2 u_1 \cdots u_3}^{(l)}(\omega_{i_l}) &= \chi_{u_{l+1} u_2 u_1 \cdots u_3}^{(l)}(\omega_n)\end{aligned}\quad (\text{E25})$$

而当 $\max\{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}\} \ll \omega_0$ 条件成立时，介质肯定同时也是无吸收的（参与非线性过程的入射光能量均太低，不足以引起介质的共振吸收），因此在该条件下③.中的全置换对称性 B 即式(E24)成立。

联立式(E24)以及式(E25)，再根据 $\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}$ 与 $\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}$ 无关，即有：

$$\begin{aligned}\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n) &= \chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n; \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}) = \chi_{u_3 u_2 u_1 \cdots u_{l+1}}^{(l)}(\omega_{i_2}; -\omega_{i_1}, \omega_n, \dots, -\omega_{i_l}) \\ &= \chi_{u_3 u_2 u_1 \cdots u_{l+1}}^{(l)}(\omega_{i_2}) = \chi_{u_3 u_2 u_1 \cdots u_{l+1}}^{(l)}(\omega_n)\end{aligned}\quad (\text{E26})$$

同理，还有 $\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n) = \chi_{u_{l+1} u_2 u_3 \cdots u_1}^{(l)}(\omega_n)$ 、 $\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n) = \chi_{u_{l+1} u_2 u_1 \cdots u_3}^{(l)}(\omega_n)$ 等，则有：

$$\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)}(\omega_n) = \chi_{u_3 u_2 u_1 \cdots u_{l+1}}^{(l)}(\omega_n) = \chi_{u_{l+1} u_2 u_3 \cdots u_1}^{(l)}(\omega_n) = \chi_{u_{l+1} u_2 u_1 \cdots u_3}^{(l)}(\omega_n) = \dots \quad (\text{E27})$$

以至于有：

$$\chi_{u_1 u_2 u_3 \cdots u_{l+1}}^{(l)} = \chi_{u_3 u_2 u_1 \cdots u_{l+1}}^{(l)} = \chi_{u_{l+1} u_2 u_3 \cdots u_1}^{(l)} = \chi_{u_{l+1} u_2 u_1 \cdots u_3}^{(l)} = \dots \quad (\text{E28})$$

此即克莱曼对称性，在非线性系数的色散可以被忽略^①的介质或场景中恒成立。

而即使在色散无法忽略的色散介质中，如果 $\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_l}$ 中有一些相等，那么相等的这些频率及其角标也可以同时两两互换，但这些频率不能与其他与其不等的频率之间连同相应角标一起互换，所以无法跨频率守恒方程的等号地调换频率和角标；且此时(E28)中不能省略频率因变量。比如倍频过程，可有：

$$\begin{aligned}\chi_{u_1 u_2 u_3}^{(l)}(\omega_n; \frac{1}{2}\omega_n, \frac{1}{2}\omega_n) &= \chi_{u_1 u_3 u_2}^{(l)}(\omega_n, \frac{1}{2}\omega_n, \frac{1}{2}\omega_n) \neq \chi_{u_3 u_1 u_2}^{(l)}(\frac{1}{2}\omega_n; \omega_n, -\frac{1}{2}\omega_n) \\ \chi_{u_1 u_2 u_3}^{(l)}(\omega_n) &= \chi_{u_1 u_3 u_2}^{(l)}(\omega_n) \neq \chi_{u_3 u_1 u_2}^{(l)}(\frac{1}{2}\omega_n) = \chi_{u_2 u_1 u_3}^{(l)}(\frac{1}{2}\omega_n) \\ &\neq \chi_{u_3 u_2 u_1}^{(l)}(\frac{1}{2}\omega_n) = \chi_{u_2 u_3 u_1}^{(l)}(\frac{1}{2}\omega_n)\end{aligned}\quad (\text{E29})$$

附E4.1 当克莱曼对称性成立时，引入的缩写记法

由于历史的原因，当克莱曼对称性成立时，对二阶张量引入了以下记法：

$$d_{u_1 u_2 u_3} := \frac{1}{2} \chi_{u_1 u_2 u_3}^{(2)} \quad (\text{E30})$$

根据式(E2)，克莱曼对称性成立时，二阶和频 ($\omega_{m_1} \neq \omega_{m_2}$) 过程生成的极化光为：

$$\tilde{P}_{u_1}^{(2)}(\omega_n) = \epsilon_0 \sum_{\omega_n = \omega_{i_1} + \omega_{i_2}} D_{i_1, i_2} \sum_{u_2 u_3}^{\{i_1, i_2, i_3\}} \chi_{u_1 u_2 u_3}^{(2)}(\omega_n; \omega_{i_1}, \omega_{i_2}) \tilde{E}_{u_2}(\omega_{i_1}) \tilde{E}_{u_3}(\omega_{i_2})$$

^① 这句话用得非常精妙，boyd 并没有说“在无色散介质中...”，而是包括了无色散介质和有色散但色散可以忽略的介质。当然，一般来说现实生活中所有介质均是色散介质，没有绝对的无色散介质。但另一方面，所有介质在甚低频的入射混频光下也全都近似于无色散介质。而无色散介质一定是无吸收介质，但无吸收介质不一定是无色散介质，前者是后者的充分必要条件，正如“克莱曼对称性成立”是“全置换对称性成立”的充分必要条件一样，前者约束更多结论却也更简单，而后者适用范围更广。

$$= \varepsilon_0 \sum_{\omega_n = \omega_{l_1} + \omega_{l_2}} 2 \sum_{u_2 u_3}^{\{i,j,k\}} \chi_{u_1 u_2 u_3}^{(2)} \tilde{E}_{u_2}(\omega_{l_1}) \tilde{E}_{u_3}(\omega_{l_2}) = \varepsilon_0 \sum_{\omega_n = \omega_{l_1} + \omega_{l_2}} \sum_{u_2 u_3}^{\{i,j,k\}} 4 d_{u_1 u_2 u_3} \tilde{E}_{u_2}(\omega_{l_1}) \tilde{E}_{u_3}(\omega_{l_2}) \quad (\text{E31})$$

去掉两边的时间部分，得到：

$$P_{u_1}^{(2)}(\omega_n) = \varepsilon_0 \sum_{\omega_n = \omega_{l_1} + \omega_{l_2}} \sum_{u_2 u_3}^{\{i,j,k\}} 4 d_{u_1 u_2 u_3} \tilde{E}_{u_2}(\omega_{l_1}) \tilde{E}_{u_3}(\omega_{l_2}) \quad (\text{E32})$$

根据克莱曼对称性，首先 $\chi_{u_1 u_2 u_3}^{(2)}$ 与 $d_{u_1 u_2 u_3}$ 的后两个角标可以互换，因此将原本九种 $u_2 u_3$ 的排列 11、22、33、23,32、13,31、12,21^①，分别记为新单独下角标 u_{23} 的六个取值 1、2、3、4、5、6，此时 $d_{u_1 u_2 u_3}$ 因与 $d_{u_1 u_{23}}$ 一一对应而等价，记号 $d_{u_1 u_{23}}$ 相比 $d_{u_1 u_2 u_3}$ 便可以少 1/3 的分量来描述 $d_{u_1 u_2 u_3}$ ，如 d_{25} 可同时代表 d_{213} 和 d_{231} 的值；并且非线性系数从原三维三阶张量 $d_{u_1 u_2 u_3}$ 降了一阶为 3×6 的二阶张量 $d_{u_1 u_{23}}$ ：

$$d_{u_1 u_{23}} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix} \quad (\text{E33})$$

而又由于克莱曼对称性， $d_{u_1 u_2 u_3}$ 的一三角标 u_1, u_3 之间和一二角标 u_1, u_2 之间，也可互换，因此记号 $d_{u_1 u_{23}}$ 的 18 个分量也不是全都两两独立的，其中有 8 对分量两两相关，以至于只剩 10 个分量两两独立：

$$\begin{aligned} \color{red}{d_{12}} &= d_{122} = d_{212} = d_{221} = \color{red}{d_{26}} \\ \color{green}{d_{13}} &= d_{133} = d_{313} = d_{331} = \color{green}{d_{35}} \\ \color{blue}{d_{14}} &= d_{123} = d_{213} = d_{231} = \color{blue}{d_{25}} \\ &= d_{312} = d_{321} = \color{blue}{d_{36}} \\ \color{brown}{d_{15}} &= d_{113} = d_{311} = \color{brown}{d_{31}} \\ \color{purple}{d_{16}} &= d_{112} = d_{211} = \color{purple}{d_{21}} \\ \color{teal}{d_{23}} &= d_{233} = d_{323} = d_{332} = \color{teal}{d_{34}} \\ \color{red}{d_{24}} &= d_{223} = d_{322} = \color{red}{d_{32}} \end{aligned} \quad (\text{E34})$$

将等号前的角标值较小的系数，替换掉等号后角标较大的系数， $d_{u_1 u_{23}}$ 可化简为：

$$d_{u_1 u_{23}} = \begin{bmatrix} d_{11} & \color{red}{d_{12}} & \color{green}{d_{13}} & \color{blue}{d_{14}} & \color{brown}{d_{15}} & \color{purple}{d_{16}} \\ \color{purple}{d_{16}} & d_{22} & \color{teal}{d_{23}} & \color{red}{d_{24}} & \color{blue}{d_{14}} & \color{red}{d_{12}} \\ \color{brown}{d_{15}} & \color{red}{d_{24}} & d_{33} & \color{teal}{d_{23}} & \color{blue}{d_{13}} & \color{red}{d_{14}} \end{bmatrix} \quad (\text{E35})$$

于是式(E32)的复振幅（定态/不含时）表达式可写为：

^① 这里的 1,2,3 分别表示 i,j,k，即 x,y,z(分量)。

$$\begin{bmatrix} P_1^{(2)}(\omega_n) \\ P_2^{(2)}(\omega_n) \\ P_3^{(2)}(\omega_n) \end{bmatrix} = 4\epsilon_0 \sum_{\omega_n = \omega_{i_1} + \omega_{i_2}} \begin{bmatrix} d_{11} \textcolor{red}{d}_{12} \textcolor{green}{d}_{13} \textcolor{blue}{d}_{14} \textcolor{brown}{d}_{15} \textcolor{purple}{d}_{16} \\ \textcolor{violet}{d}_{16} d_{22} \textcolor{blue}{d}_{23} \textcolor{red}{d}_{24} \textcolor{blue}{d}_{14} \textcolor{red}{d}_{12} \\ \textcolor{brown}{d}_{15} \textcolor{red}{d}_{24} d_{33} \textcolor{blue}{d}_{23} \textcolor{green}{d}_{13} \textcolor{blue}{d}_{14} \end{bmatrix} \cdot \begin{bmatrix} E_1(\omega_{i_1})E_1(\omega_{i_2}) \\ E_2(\omega_{i_1})E_2(\omega_{i_2}) \\ E_3(\omega_{i_1})E_3(\omega_{i_2}) \\ E_2(\omega_{i_1})E_3(\omega_{i_2}) + E_3(\omega_{i_1})E_2(\omega_{i_2}) \\ E_1(\omega_{i_1})E_3(\omega_{i_2}) + E_3(\omega_{i_1})E_1(\omega_{i_2}) \\ E_1(\omega_{i_1})E_2(\omega_{i_2}) + E_2(\omega_{i_1})E_1(\omega_{i_2}) \end{bmatrix} \quad (\text{E36})$$

考虑其中给定的一对的混频光 i_1, i_2 所产生的和频极化光，有：

$$\begin{bmatrix} P_x^{(2)}(\omega_{i_1} + \omega_{i_2}) \\ P_y^{(2)}(\omega_{i_1} + \omega_{i_2}) \\ P_z^{(2)}(\omega_{i_1} + \omega_{i_2}) \end{bmatrix} = 4\epsilon_0 \begin{bmatrix} d_{11} \textcolor{red}{d}_{12} \textcolor{green}{d}_{13} \textcolor{blue}{d}_{14} \textcolor{brown}{d}_{15} \textcolor{purple}{d}_{16} \\ \textcolor{violet}{d}_{16} d_{22} \textcolor{blue}{d}_{23} \textcolor{red}{d}_{24} \textcolor{blue}{d}_{14} \textcolor{red}{d}_{12} \\ \textcolor{brown}{d}_{15} \textcolor{red}{d}_{24} d_{33} \textcolor{blue}{d}_{23} \textcolor{green}{d}_{13} \textcolor{blue}{d}_{14} \end{bmatrix} \cdot \begin{bmatrix} E_x(\omega_{i_1})E_x(\omega_{i_2}) \\ E_y(\omega_{i_1})E_y(\omega_{i_2}) \\ E_z(\omega_{i_1})E_z(\omega_{i_2}) \\ E_y(\omega_{i_1})E_z(\omega_{i_2}) + E_z(\omega_{i_1})E_y(\omega_{i_2}) \\ E_x(\omega_{i_1})E_z(\omega_{i_2}) + E_z(\omega_{i_1})E_x(\omega_{i_2}) \\ E_x(\omega_{i_1})E_y(\omega_{i_2}) + E_y(\omega_{i_1})E_x(\omega_{i_2}) \end{bmatrix} \quad (\text{E37})$$

对于产生特定频率二次谐波(SHG)的倍频极化光，此时 $i_1 = i_2$ 代表同一束参与混频的单色光，相应非线性光学过程为：

$$\begin{bmatrix} P_x^{(2)}(2\omega_{i_1}) \\ P_y^{(2)}(2\omega_{i_1}) \\ P_z^{(2)}(2\omega_{i_1}) \end{bmatrix} = 2\epsilon_0 \begin{bmatrix} d_{11} \textcolor{red}{d}_{12} \textcolor{green}{d}_{13} \textcolor{blue}{d}_{14} \textcolor{brown}{d}_{15} \textcolor{purple}{d}_{16} \\ \textcolor{violet}{d}_{16} d_{22} \textcolor{blue}{d}_{23} \textcolor{red}{d}_{24} \textcolor{blue}{d}_{14} \textcolor{red}{d}_{12} \\ \textcolor{brown}{d}_{15} \textcolor{red}{d}_{24} d_{33} \textcolor{blue}{d}_{23} \textcolor{green}{d}_{13} \textcolor{blue}{d}_{14} \end{bmatrix} \cdot \begin{bmatrix} E_x(\omega_{i_1})E_x(\omega_{i_1}) \\ E_y(\omega_{i_1})E_y(\omega_{i_1}) \\ E_z(\omega_{i_1})E_z(\omega_{i_1}) \\ 2E_y(\omega_{i_1})E_z(\omega_{i_1}) \\ 2E_x(\omega_{i_1})E_z(\omega_{i_1}) \\ 2E_x(\omega_{i_1})E_y(\omega_{i_1}) \end{bmatrix} \quad (\text{E38})$$