第三届阿里巴巴全球数学竞赛预赛部分参考答案

1,2. 在一个虚拟的世界中, 每个居民(设想为没有大小的几何点) 依次编号为 $1, 2, \cdots$. 为了抗击某种疫情, 这些居民要接种某疫苗，并在注射后在现场留观一段时间。现在假设留观的场所是平面上的一个半径为 $\frac{1}{4}$ 的圆周。为了安全，要求第 m 号居民和第 n 号居民之间的距离 $d_{m,n}$ 满足

$$(m+n)d_{m,n} \geq 1,$$

这里我们考虑的是圆周上的距离, 也就是两点间劣弧的弧长。那么

1 **选择题(4分)** 下列选项( ) 符合实际情况。

A 这个留观室最多能容纳 8 个居民;

B 这个留观室能容纳的居民个数有大于 8 的上限;

C 这个留观室可以容纳任意多个居民。

2 **证明题(6分)** 证明你的论断。

R1 答案. 选项C 符合实际情况.

R2 答案. 解法一. 我们可以按下述方式安排第 $1, 2, \ldots$ 号居民的位置. 首先, 任意安排第 1 号居民的位置。对 $n \geq 2$, 若第 $1, 2, \ldots, n-1$ 号居民的位置已经被安排好, 我们考虑第 n 号居民不能在哪些位置。

对于 $1 \leq m \leq n-1$, 由 $d_{m,n} \geq \frac{1}{m+n}$, 我们知道, 从第 m 号居民的位置开始, 沿顺、逆时针方向各走 $\frac{1}{m+n}$ 的距离, 所形成的长度为 $\frac{2}{m+n}$ 的圆弧内部是不可以安排第 n 号居民的. 而这些圆弧的总长度

$$\frac{2}{n+1}+\frac{2}{n+2}+\cdots+\frac{2}{2n-1} < 2(\ln\frac{n+1}{n}+\ln\frac{n+2}{n+1}+\cdots+\ln\frac{2n-1}{2n-2}) = 2\ln\frac{2n-1}{n} < 2\ln 2.$$

因此, 这些圆弧的并集的总长度不超过 $2\ln 2$, 而整个圆周长为 $\frac{1}{4}\cdot 2\pi = \frac{\pi}{2}$. 熟知 $\frac{\pi}{2} > 1.5 > 2\ln 2$, 故这些圆弧不能覆盖整个圆周, 因此第 n 号居民总可以选择一个合适的位置, 使得他与第 $1, 2, \ldots, n-1$ 号居民之间的距离均满足题目条件. 由数学归纳法可知，这个圆周可以容纳任意多个居民.

解法二. 我们以圆周的圆心为原点建立平面直角坐标系，并将第 $1, 2, 3, 4$ 号居民分别放在 $(\frac{1}{4}, 0), (-\frac{1}{4}, 0), (0, \frac{1}{4}), (0, -\frac{1}{4})$ 处, 即他们的辐角主值分别为 $0, \pi, \frac{\pi}{2}, \frac{3\pi}{2}$. 此时任意两名居民的距离不小于 $2\cdot\frac{1}{4}\cdot\frac{\pi}{4} = \frac{\pi}{8} > \frac{1}{1+2}$, 故此时的 4 名居民满足题目条件.

我们使用数学归纳法证明下面命题: 对整数 $k > 2$, 可以将第 $1, 2, \ldots, 2^k$ 号居民安置于圆周上一个内接正 2^k 边形的各个顶点处, 使得它们互相之间(在圆周上的) 距离满足题目条件, 且编号为 $1, 2, \ldots, 2^{k-1}$ 号的居民在圆周上两两不相邻.

上述命题对 $k = 2$ 成立. 若其对 k 成立，即前 2^k 号居民的位置都已确定. 考虑他们将圆周分成的 2^k 段弧. 我们要将第 $2^k+1, 2^k+2, \ldots, 2^{k+1}$ 号居民放置在这些弧的中点. 现在来证明可以适当放置使得涉及第 $2^k+1, 2^k+2, \ldots, 2^{k+1}$ 号居民的距离均满足题目条件.

我们将第 $2^k+1, 2^k+2$ 号居民放置在与第 2^{k-1} 号居民相邻的位置(即与第 2^{k-1} 号居民的辐角差为 $\frac{2\pi}{2^{k+1}}$ 距离的位置); 将第 $2^k+3, 2^k+4$ 号居民放置在与第 $2^{k-1}-1$ 号居民相邻的位置; $\cdots$ 将第 $2^k+2a-1, 2^k+2a$ 号居民放置在与第 $2^{k-1}-a+1$ 号居民相邻的位置; $\cdots$ 将第 $2^{k+1}-1, 2^{k+1}$ 号居民放置在与第 1 号居民相邻的位置.

由于前 2^{k-1} 号居民在圆周上两两不相邻, 这样的放置是可行的. 现在考虑任意两名居民的距离(只需考虑至少一位居民是“新”的情形). 因为圆周被分成了 2^{k+1} 段, 每段弧长为 $\frac{2\pi}{2^{k+1}}\cdot\frac{1}{4} = \frac{\pi}{2^{k+2}}$, 对于两位编号分别为 $m > 2^k$ 和 n 的居民, 若它们之间至少有两段弧, 则

$$d_{m,n} \geq \frac{\pi}{2^{k+1}} > \frac{\pi}{2}\cdot\frac{1}{m+n}\cdot\frac{2(m+n)}{2^{k+1}} > \frac{1}{m+n};$$

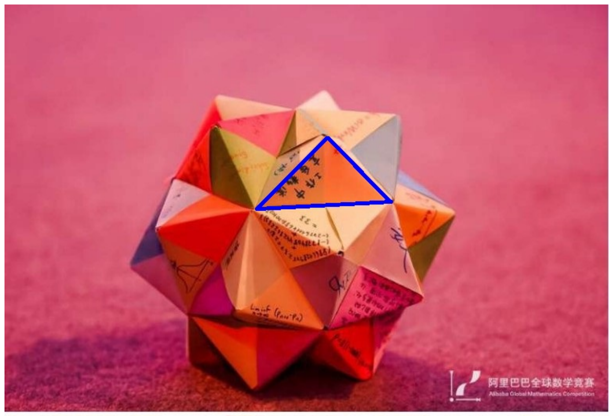
若他们之间的距离恰为一段弧长，设 $n \in \{2^k+2a-1, 2^k+2a\}$, 则 $m > 2^{k-1}-a+1$, 因

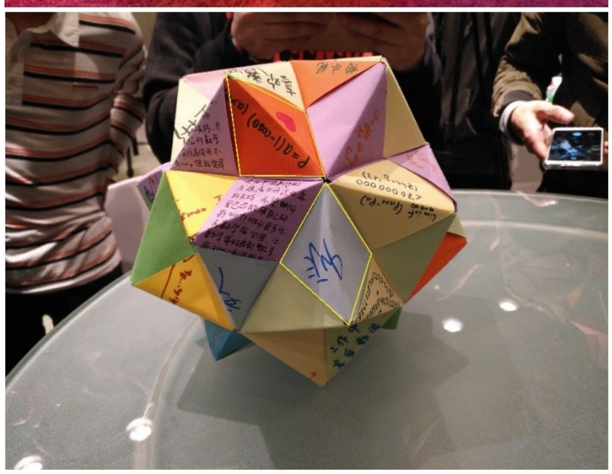
此

$$(m+n)d_{m,n} \geq \frac{\pi}{2^{k+2}} \cdot (2^k + 2a - 1 + 2^{k-1} - a + 1) = \frac{\pi}{2^{k+2}} \cdot (2 \cdot 2^{k-1} + a) \geq \frac{3\pi \cdot 2^{k-1}}{2^{k+1}} = \frac{3\pi}{8} > 1.$$

所以，第 $1, 2, \ldots, 2^{k+1}$ 号居民两两之间的距离均满足题目条件. 由数学归纳法知, 可以安排任意多名居民.

3,4. 2019年第一届阿里巴巴数学竞赛的优胜者们在参加集训营的时候, 集体送给主办方负责人的礼物, 是一个有 60 个全等的三角形面的多面体。从图中我们可以看到, 这个多面体的表面是 60 个全等的空间四边形拼接而成的。





一个空间 n 边形是指由一个平面 n 边形沿若干条对角线做适当翻折(即在选定的对角线处形成适当的二面角) 后得到的空间图形。两个空间图形全等指的是它们可以通过 $\mathbb{R}^3$ 中的一个等距变换完全重合。一个多面体指的是一个空间有界区域, 其边界可以由有限多个平面多边形沿公共边拼接而成。

3 **判断题(4分)** 我们知道 $2021 = 43 \times 47$. 那么是否存在一个多面体, 它的表面可以由 43 个全等的空间 47 边形拼接而成?

4 **问答题(6分)** 请对你的判断给出逻辑的解释。

R3 答案. 可以.

R4 答案. 我们只需要举一个例子即可. 考虑一个标准的环面 $\mathbb{T}$, 其上的点可以由两个参数来表示:

$$\mathbb{T} = \{\theta, \varphi :\ 0 \le \theta, \varphi < 2\pi\}.$$

我们可以认为这个环面以 z-轴为对称轴: (θ, φ) 对应于空间中的 $((R + r\cos\varphi)\cos\theta, (R + r\cos\varphi)\sin\theta, r\sin\varphi)$.

对于 $1 \le k \le 43$, 我们考虑环面上的区域

$$D_k = \{\theta, \varphi :\ \frac{2(k-1)}{43}\pi + 3\frac{\varphi}{86} \le \theta \le \frac{2k}{43}\pi + 3\frac{\varphi}{86}\}.$$

直观地说, 把环面分成全等的 43 份之后, 每一份沿 $\{\varphi = 0\}$ 切开, 将切开处的一侧保持不动, 另一侧扭转一定角度.

现在, 把 $\{\varphi = 0\}$ 这个圆变形成一个正 43 边形, 各个顶点分别对应于 $\theta = \frac{2k}{43}\pi$. 这样 D_k 有四条"边"(其中两条位于 $\{\varphi = 0\}$ 上), 四个顶点(两个位于正 43 边形的顶点处, 两个位于边的中点处, 我们还要标记出这两个中点之间的正 43 边形的顶点). 记为

$$\begin{aligned}
&C_{k,0} = (\frac{2(k-1)}{43}\pi, 0), C_{k,1} = (\frac{2k}{43}\pi, 0);\\
&D_{k,0} = (\frac{2k+1}{43}\pi, 2\pi), D_{k,1} = (\frac{2k+3}{43}\pi, 2\pi);\\
&E_k = (\frac{2k+2}{43}\pi, 2\pi).
\end{aligned}$$

在 ∂D_k 的另一条边上取 21 个点, 例如

$$A_{k,i} = (\frac{2(k-1)}{43}\pi + 3\frac{\varphi}{86}\pi, \frac{i}{11}\pi), i = 1, \dots, 21.$$

然后绕 z-轴旋转 $\frac{2}{43}\pi$ 后得到另 21 个点, 记为 $B_{k,i}, i = 1, \dots, 21$.

连结线段 $C_{k,0}C_{k,1}, C_{k,0}A_{k,1}, C_{k,1}B_{k,1}, A_{k,i}A_{k,i+1}, B_{k,i}B_{k,i+1}, A_{k,i}B_{k,i}, A_{k,i}B_{k,i+1}$ $(i = 1, \dots, 21)$, 以及 $A_{k,21}D_{k,0}, B_{k,21}D_{k,1}, A_{k,21}E_k, B_{k,21}E_k, D_{k,0}E_k$ 和 $E_kD_{k,1}$. 我们得到一个空间 47 边形. 这样我们就得到了 43 个全等的(上述构造与 k 无关) 空间 47 边形, 它们能够拼出一个多面体.

说明: 一个典型的错误是误认为这些空间多边形的顶点(边) 都是多面体的顶点(边), 从而根据"每条边都算两次" 和" 2021 是奇数" 得到"矛盾", 由此认为本题的解答是否定的.

5. 去年，张师傅因为多旋圈面爆红，今年他来到了达摩院给扫地僧做面。某天，软件工程师小李跟张师傅吐槽工作。小李主要研究和设计算法用于调节各种产品的参数。这样的参数一般可以通过极小化 $\mathbb{R}^n$ 上的某个损失函数 f 求得。在小李最近的一个项目中，这个损失函数是另外一个课题组提供的；出于安全考虑和技术原因，该课题组难以向小李给出此函数的内部细节，而只能提供一个接口用于计算任意 $\mathbf{x} \in \mathbb{R}^n$ 处的函数值 $f(\mathbf{x})$。所以，小李必须仅基于函数值来极小化 f。而且，每次计算 f 的值都消耗不小的计算资源。好在该问题的维度 n 不是很高 (10 左右)。另外，提供函数的同事还告知小李不妨先假设 f 是光滑的。

这个问题让张师傅想起了自己收藏的一台古董收音机。要在这台收音机上收听一个节目，你需要小心地来回拧一个调频旋钮，同时注意收音效果，直到达到最佳。在这过程中，没有人确切地知道旋钮的角度和收音效果之间的定量关系是什么。张师傅和小李意识到，极小化 f 不过就是调节一台有多个旋钮的机器: 想象 $\mathbf{x}$ 的每一个分量由一个旋钮控制，而 $f(\mathbf{x})$ 表示这台机器的某种性能，只要我们来回调整每个旋钮，同时监视 f 的值，应该就有希望找到最佳的 $\mathbf{x}$。受此启发，两人一起提出了极小化 f 的一个迭代算法，并命名为"自动前后调整算法" (Automated Forward/Backward Tuning，AFBT，算法 1)。在第 k 次迭代中，AFBT 通过前后调整 $\mathbf{x}_k$ 的单个分量得到 $2n$ 个点 $\{\mathbf{x}_k \pm t_k\mathbf{e}^i : i = 1, \ldots, n\}$, 其中 t_k 为步长；然后，令 $\mathbf{y}_k$ 为这些点中函数值最小的一个，并检查 $\mathbf{y}_k$ 是否使 f 充分减小；若是，取 $\mathbf{x}_{k+1} = \mathbf{y}_k$，并将步长增倍；否则，令 $\mathbf{x}_{k+1} = \mathbf{x}_k$ 并将步长减半。在算法 1 中， $\mathbf{e}^i$ 表示 $\mathbb{R}^n$ 中的第 i 个坐标向量，它的第 i 个分量为 1, 其余皆为 0；$\mathbb{1}(\cdot)$ 为指示函数 —— 若 $f(\mathbf{x}_k) - f(\mathbf{y}_k)$ 至少为 t_k 之平方，则 $\mathbb{1}[f(\mathbf{x}_k) - f(\mathbf{y}_k) \geq t_k^2]$ 取值为 1，否则为 0。

算法 1 自动前后调整算法 (AFBT)

输入 $\mathbf{x}_0 \in \mathbb{R}^n$, $t_0 > 0$。对 $k = 0, 1, 2, \ldots,$ 执行以下循环。

1: $\mathbf{y}_k := \operatorname{argmin}\{f(\mathbf{y}) : \mathbf{y} = \mathbf{x}_k \pm t_k\mathbf{e}^i,\ i = 1, \ldots, n\}$ # 计算损失函数。
2: $s_k := \mathbb{1}[f(\mathbf{x}_k) - f(\mathbf{y}_k) \geq t_k^2]$ # 是否充分下降？ 是：$s_k = 1$； 否：$s_k = 0$。
3: $\mathbf{x}_{k+1} := (1 - s_k)\mathbf{x}_k + s_k\mathbf{y}_k$ # 更新迭代点。
4: $t_{k+1} := 2^{2s_k - 1}t_k$ # 更新步长。$s_k = 1$：步长增倍；$s_k = 0$：步长减半。

现在，我们对损失函数 $f : \mathbb{R}^n \to \mathbb{R}$ 作出如下假设。

假设 1. f 为凸函数，即对任何 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 与 $\alpha \in [0, 1]$ 都有

$$f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) \ \leq\ (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y}).$$

假设 2. f 在 $\mathbb{R}^n$ 上可微且 ∇f 在 $\mathbb{R}^n$ 上 L-Lipschitz 连续。

假设 3. f 的水平集有界，即对任意 $\lambda \in \mathbb{R}$，集合$\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq \lambda\}$ 皆有界。

基于**假设 1** 与**假设 2**，可以证明

$$\langle \nabla f(\mathbf{x}),\ \mathbf{y}-\mathbf{x} \rangle \ \le\ f(\mathbf{y})-f(\mathbf{x})\ \le\ \langle \nabla f(\mathbf{x}),\ \mathbf{y}-\mathbf{x} \rangle + \frac{L}{2}\|\mathbf{x}-\mathbf{y}\|^2$$

对任何 $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ 成立；**假设 1** 与**假设 3** 则保证f 在 $\mathbb{R}^n$ 上取到有限的最小值 f^*。凸函数的更多性质可参考任何一本凸分析教科书。

证明题(20分) 在**假设 1–3**下，对于 AFBT，证明

$$\lim_{k\to\infty} f(\mathbf{x}_k)\ =\ f^*.$$

R5 证明. 假设 $f(\mathbf{x}_k) \nrightarrow f^*$。因 $\{f(\mathbf{x}_k)\}$ 不增, 故 $\inf_{k\geq 0}[f(\mathbf{x}_k)-f^*] > 0$。记$\mathbf{g}_k = \nabla f(\mathbf{x}_k)$, 则 $\inf_{k\geq 0}\|\mathbf{g}_k\| > 0$ (取 $\mathbf{x}^*$ 使 $f(\mathbf{x}^*) = f^*$, 则 f 之凸性保证 $\langle \mathbf{g}_k, \mathbf{x}_k - \mathbf{x}^*\rangle \geq f(\mathbf{x}_k) - f^*$；同时，$\{f(\mathbf{x}_k)\}$ 之单调性与 f 之水平集有界性保证 $\{\mathbf{x}_k - \mathbf{x}^*\}$ 有界。故 $\inf_{k\geq 0}\|\mathbf{g}_k\| > 0$)。换言之，存在 $\varepsilon > 0$ 使 $\|\mathbf{g}_k\| \geq \varepsilon$ 对所有 $k \geq 0$ 成立。任给 $k \geq 0$, 可取 $i_k \in \{1, \ldots, n\}$ 满足

$$|\langle \mathbf{g}_k, \mathbf{e}^{i_k}\rangle| \ \geq \ \kappa\|\mathbf{g}_k\| \ \geq \ \kappa\varepsilon, \tag{1}$$

其中 $\kappa = 1/\sqrt{n}$。故

$$f(\mathbf{y}_k) \ \leq \ \min\{f(\mathbf{x}_k \pm t_k\mathbf{e}^{i_k})\} \ \leq \ f(\mathbf{x}_k) - t_k|\langle \mathbf{g}_k, \mathbf{e}^{i_k}\rangle| + \frac{Lt_k^2}{2} \ \leq \ f(\mathbf{x}_k) - \kappa\varepsilon t_k + \frac{Lt_k^2}{2}. \tag{2}$$

所以，只要

$$t_k \ \leq \ \frac{2\kappa\varepsilon}{L+2}, \tag{3}$$

我们就有 $f(\mathbf{y}_k) \leq f(\mathbf{x}_k) - t_k^2$, 从而 $s_k = 1$, 进而有 $t_{k+1} = 2t_k$。由此，易见

$$t_k \ \geq \ \underline{t} \ \equiv \ \min\left\{t_0, \ \frac{\kappa\varepsilon}{L+2}\right\} \ > \ 0 \tag{4}$$

对所有 $k \geq 0$ 成立。故存在无穷个 k 使得 $s_k = 1$ (否则 $t_k \to 0$); 对每一个这样的 k, $f(\mathbf{x}_k) - f(\mathbf{x}_{k+1}) \geq \underline{t}^2$。这与 f 下方有界相矛盾。故所论成立。

6. 令 n 为正整数。对任一正整数 k，记 $0_k = \mathrm{diag}\{\underbrace{0,\ldots,0}_{k}\}$ 为 $k\times k$ 的零矩阵。令

$$Y = \begin{pmatrix} 0_n & A \\ A^t & 0_{n+1} \end{pmatrix}$$

为一个 $(2n+1)\times(2n+1)$ 矩阵, 其中 $A = (x_{i,j})_{1\le i\le n, 1\le j\le n+1}$ 是一个 $n\times(n+1)$ 实矩阵且 A^t 记 A 的转置矩阵，即 $(n+1)\times n$ 的矩阵，(j,i) 处元素为 $x_{i,j}$.

(i) **证明题(10分)** 称复数 λ 为 $k\times k$ 矩阵 X 的一个特征值, 如果存在非零列向量

$$v = (x_1,\ldots,x_k)^t$$

使得 $Xv = \lambda v$. 证明: 0 是 Y 的特征值且 Y 的其他特征值形如 $\pm\sqrt{\lambda}$, 其中非负实数 λ 是 AA^t 的特征值。

(ii) **证明题(15分)** 令 $n=3$ 且 a_1,a_2,a_3,a_4 是 4 个互不相等的正实数。记

$$a = \sqrt{\sum_{1\le i\le 4} a_i^2}$$

以及

$$x_{i,j} = a_i\delta_{i,j} + a_j\delta_{4,j} - \frac{1}{a^2}(a_i^2 + a_4^2)a_j$$

$(1\le i\le 3,\ 1\le j\le 4)$, 其中 $\delta_{i,j} = \begin{cases} 1 \text{ if } i=j \\ 0 \text{ if } i\ne j \end{cases}$. 证明: Y 有 7 个互不相等的特征值。

R6 证明: (i) 记 $I_n = \mathrm{diag}\{\underbrace{1, \dots, 1}_{n}\}$ 为 $n \times n$ 恒同矩阵. 作初等变换可证

$$\det(\lambda I_{2n+1} - Y) = \lambda \det(\lambda^2 I_n - AA^t).$$

所以, 0 是 Y 的特征值且 Y 的其他特征值形如 $\pm\sqrt{\lambda}$, 其中 λ 是 AA^t 的非负实数特征值.

(ii) 记 $u = (a_4, a_4, a_4)$, $v = (\frac{a_1^2+a_4^2}{a}, \frac{a_2^2+a_4^2}{a}, \frac{a_3^2+a_4^2}{a})$. 计算得

$$AA^t = \mathrm{diag}\{a_1^2, \dots, a_3^2\} + u^t u - v^t v.$$

设 $f(s) = \det(sI_n - AA^t)$ 为 AA^t 的特征多项式. 计算得

$$f(a_i^2) = \frac{a_i^2}{a^2} \prod_{1 \le j \le 4, j \ne i} (a_i^2 - a_j^2)$$

$(\forall i \in \{1, 2, 3, 4\})$. 令 a_1', a_2', a_3', a_4' 为 a_1, a_2, a_3, a_4 经重排后得到的递减序列. 由 $f(x_i^2)$ 的表达式得: AA^t 有三个互不相等的特征值 b_1^2, b_2^2, b_3^3, 其中 b_1, b_2, b_3 是满足

$$a_1' > b_1 > a_2' > b_2 > a_3' > b_3 > a_4'$$

的正实数. 因此, 由(i)得 Y 有 7 个互不相等的特征值.

7. 对于 $\mathbb{R}$ 上的连续且绝对可积的复数值函数 $f(x)$，定义 $\mathbb{R}$ 上的函数 $(Sf)(x)$:

$$(Sf)(x) = \int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} ux} f(u) \mathrm{d}u.$$

(i) **问答题(10分)** 求 $S(\frac{1}{1+x^2})$ 和 $S(\frac{1}{(1+x^2)^2})$ 的显式表达式。

(ii) **问答题(15分)** 对任意整数 k，记 $f_k(x) = (1+x^2)^{-1-k}$. 假设 $k \geq 1$，找到常数 c_1, c_2 使得函数 $y = (Sf_k)(x)$ 满足二阶常微分方程

$$xy'' + c_1 y' + c_2 xy = 0.$$

R7 答案:

(i)

$$S(\frac{1}{1+x^2}) = \pi e^{-2\pi|x|}$$

且

$$S(\frac{1}{(1+x^2)^2}) = \frac{\pi}{2}(1+2\pi|x|)e^{-2\pi|x|}.$$

(ii) $c_1 = -2k$ 且 $c_2 = -4\pi^2$.

解答 记 V 为 $\mathbb{R}$ 上的复数值、连续、绝对可积的函数组成的线性空间.

Lemma 0.1. *(i)* 若 $f(x) \in V$, $f'(x) \in V$ 且 $\lim_{x\to\infty} f(x) = 0$，则

$$(Sf')(x) = -2\pi \mathbf{i} x(Sf)(x). \tag{5}$$

(ii) 若 $f(x) \in V$ 且 $xf(x) \in V$，则

$$(Sf)' = 2\pi \mathbf{i} S(xf(x)). \tag{6}$$

引理*0.1*的证明. (i)

$$\begin{aligned} & (Sf')(x) \\ = & \int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} ux} f'(u)\mathrm{d}u \\ = & e^{2\pi \mathbf{i} ux} f(u)|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} (e^{2\pi \mathbf{i} ux})' f(u)\mathrm{d}u \\ = & -2\pi \mathbf{i} x \int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} ux} f(u)\mathrm{d}u \\ = & -2\pi \mathbf{i} x(Sf)(x) \end{aligned}$$

(ii) 对任意的 $a, b \in \mathbb{R}$ $(a < b)$,

$$\begin{aligned} & \int_a^b 2\pi \mathbf{i} S(xf(x))\mathrm{d}x \\ = & \int_a^b 2\pi \mathbf{i} (\int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} ux} uf(u)\mathrm{d}u)\mathrm{d}x \\ = & \int_{-\infty}^{+\infty} (\int_a^b 2\pi \mathbf{i} u e^{2\pi \mathbf{i} ux} f(u)\mathrm{d}x)\mathrm{d}u \\ = & \int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} bu} f(u)\mathrm{d}u - \int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} au} f(u)\mathrm{d}u \end{aligned}$$

$$= \quad (Sf)(b) - (Sf)(a).$$

这样, $(Sf)' = 2\pi\mathbf{i}S(xf(x))$. $\square$

引理0.1有如下推论.

Corollary 0.2. *(i)* 假设 $f, f', Sf, x(Sf)(x) \in V$ 且

$$\lim_{x\to\infty} f(x) = 0.$$

若 $(S(Sf))(x) = f(-x)$, 则 $(S(Sf'))(x) = f'(-x)$.

(ii) 假设 $f(x), xf(x), Sf, (Sf)' \in V$ 且

$$\lim_{x\to\infty} (Sf)(x) = 0.$$

若 $(S(Sf))(x) = f(-x)$, 则 $S(S(xf(x))) = -xf(-x)$.

Lemma 0.3. *(i)* $S((1+x^2)^{-1}) = \pi e^{-2\pi|x|}$.

(ii) $S(\pi e^{-2\pi|x|}) = (1+x^2)^{-1}$.

Proof. (i)记 $f(x) = (1+x^2)^{-1}$. 对于 $x \geq 0$, 我们有

$$(Sf)(x) = \lim_{A\to+\infty} \int_{-A}^{A} \frac{e^{2\pi\mathbf{i}ux}}{1+u^2} \mathrm{d}u.$$

记

$$C_A := \{z = u + \mathbf{i}v : -A \leq u \leq A, v = 0\} \bigcup \{z = Ae^{\mathbf{i}\theta} : 0 \leq \theta \leq \pi\}.$$

注意到当 $A > 1$ 时, $\mathbf{i}$ 是 $\frac{1}{1+z^2}$ 在 C_A 界定的有界区域内的唯一极点. 由回路积分的方法并令 $A \to \infty$, 我们得到 $(Sf)(x) = \pi e^{-2\pi x}$. 由于 $f(x)$ 是偶函数, 所以 $(Sf)(x)$ 也是偶函数. 这样, $(Sf)(x) = \pi e^{-2\pi|x|}$.

(ii)记 $g(x) = \pi e^{-2\pi|x|}$. 直接计算得

$$\begin{aligned}
& (Sg)(x) \\
= \quad & \int_{-\infty}^{\infty} e^{2\pi\mathbf{i}xu} \pi e^{-2\pi|u|} \mathrm{d}u \\
= \quad & \pi \int_{0}^{\infty} (e^{2\pi\mathbf{i}xu} + e^{-2\pi\mathbf{i}xu}) e^{-2\pi u} \mathrm{d}u \\
= \quad & -\frac{1}{2}\left(\frac{e^{-2\pi(1+\mathbf{i}x)u}}{1+\mathbf{i}x} + \frac{e^{-2\pi(1-\mathbf{i}x)u}}{1-\mathbf{i}x}\right)\Big|_0^\infty \\
= \quad & \frac{1}{1+x^2}.
\end{aligned}$$

$\square$

Lemma 0.4. *(i)* 对于任意的 $k \geq 0$, Sf_k 形如 $(Sf_k)x = e^{-2\pi|x|}g_k(|x|)$, 其中 g_k 是个 k 次多项式.

(ii) 对于任意的 $k \geq 0$, $S(S(f_k)) = f_k$ 且 $S(S(xf_{k+1}(x))) = -xf_{k+1}(x)$.

Proof. (i)我们有递归公式

$$f_{k+1} = f_k + \frac{1}{2(k+1)}xf_k'(x).$$

由引理0.1, 得递归公式:

$$Sf_{k+1} = Sf_k - \frac{1}{2(k+1)}(x(Sf_k)(x))'.$$

由此, 由归纳法我们导出结论.

(ii)注意到 $f_k'(x) = -2(k+1)xf_{k+1}(x)$ 且 $(xf_k(x))' = -(1+2k)f_k(x)+2(k+1)f_{k+1}(x)$. 由(i)部分的结论结合引理0.1, 我们知道推论0.2中的假设对函数 $f_k(x)$ 和 $xf_{k+1}(x)(k \geq 0)$ 成立. 这样, 由归纳法可证 $S(S(f_k)) = f_k(x)$ 且 $S(S(xf_{k+1}(x))) = -xf_{k+1}(x)(k \geq 0)$. $\square$

回到题目本身. (i)在引理0.3中已经证明 $S((1+x^2)^{-1}) = \pi e^{-2\pi|x|}$. 由(5)得

$$S(-2x(1+x^2)^{-2}) = -2\pi^2\mathbf{i}xe^{-2\pi|x|}.$$

再由(6)得

$$S(-2x^2(1+x^2)^{-2}) = -\pi(1-2\pi|x|)e^{-2\pi|x|}.$$

这样,

$$S((1+x^2)^{-2}) = \frac{\pi}{2}(1+2\pi|x|)e^{-2\pi|x|}.$$

(ii)首先, 当 $k \geq 1$ 时, $x^jf_k(x)$ $(0 \leq j \leq 2k)$ 都是绝对可积的. 这样, 由(6), $y = (Sf_k)(x)$ 是 $2k$ 次连续可微函数. 由引理0.1和引理0.4得: $xy'' + c_1y' + c_2xy = 0$ 等价于

$$(x^2f_k' + 2xf_k) - c_1xf_k - \frac{c_2}{4\pi^2}f_k' = 0.$$

输入 $f_k(x) = (1+x^2)^{-1-k}$, 我们得到 $c_1 = -2k$ 且 $c_2 = -4\pi^2$.

8. 当某公司推出一个新的社交软件时，公司的市场部门除了会关心该软件的活跃客户的总人数随时间的变化，也会对客户群体的一些特征做具体的调研和分析。我们用 $n(t,x)$ 表示客户的数量密度（以下简称密度），这里 t 表示时间，而 x 表示客户对该社交软件的使用时长，那么在 t 时刻，对于 $0 < x_1 < x_2$，使用时长介于 x_1 和 x_2 之间的客户数量为 $\int_{x_1}^{x_2} n(t,x)dx$。我们假设，密度 $n(t,x)$ 随着时间演化受以下几个因素的影响：

假设 1. 当客户持续使用该社交软件时，他的使用时长随时间线性增长。

假设 2. 客户在使用过程中，可能会停止使用，我们假设停止速率 $d(x) > 0$ 只跟使用时长 x 有关。

假设 3. 新客户的来源有两个。

① 公司的宣传：单位时间内因此增加的人数是时间的函数，用 $c(t)$ 表示。

② 老客户的宣传：老客户会主动向自己的同事、朋友等推荐使用该社交软件，推荐成功的速率跟客户的使用时长 x 有关，记作 $b(x)$。

假设如果在某一时刻, 记为 $t = 0$ 时，密度函数是已知的，$n(0,x) = n_0(x)$。可以推导出，$n(t,x)$ 的时间演化满足如下的方程

$$\begin{cases} \frac{\partial}{\partial t}n(t,x) + \frac{\partial}{\partial x}n(t,x) + d(x)n(t,x) = 0, \quad t \geq 0, x \geq 0, \\ N(t) := n(t, x = 0) = c(t) + \int_0^\infty b(y)n(t,y)dy. \end{cases} \tag{7}$$

这里 $N(t)$ 可解读为新客户的增加速率。我们假设 $b, d \in L_+^\infty(0,\infty)$，即 $b(x)$ 和 $d(x)$ 正且(本质)有界。以下，我们先做一个简化假设： $c(t) \equiv 0$，即新客户的增加只跟老客户的宣传有关。

(i) **问答题(10分)** 根据**假设 1**和**假设 2**，形式地推导出(7) 中 $n(t,x)$ 所满足的偏微分方程，需要在推导过程中指出模型假设和数学表达式之间的对应关系。再根据**假设 3**，解释(7) 中 $N(t)$ 的定义的含义。

(ii) **问答题(10分)** 我们想要研究新客户的增加速率 $N(t)$ 和推荐成功速率 $b(x)$ 之间的关系。为此，请推导出一个 $N(t)$ 所满足的方程，且方程中只包含 $N(t)$, $n_0(x)$, $b(x)$, $d(x)$，而不包含 $n(t,x)$。并证明， $N(t)$ 满足如下估计

$$|N(t)| \leq \|b\|_\infty e^{\|b\|_\infty t} \int_0^\infty |n_0(x)|\, dx, \tag{8}$$

这里 $\|\cdot\|_\infty$ 表示 L^∞ 范数。

(iii) **证明题(10分)** 最后，我们想要研究，在充分长的时间之后，数量密度函数 $n(t,x)$ 有什么渐近的趋势。由于客户总人数可能一直在增加，所以我们不方便直接研究数量密度函数 $n(t,x)$，而更应该去看一个重整化的的密度函数。

为此，我们首先假设如下的特征值问题有唯一解 $(\lambda_0, \varphi(x))$：

$$\begin{cases} \varphi'(x) + (\lambda_0 + d(x))\,\varphi(x) = 0, \quad x \geq 0, \\ \varphi(x) > 0, \quad \varphi(0) = \int_0^\infty b(x)\varphi(x)dx = 1, \end{cases}$$

并且它的对偶问题也有唯一的解 $\psi(x)$:

$$\begin{cases} -\psi'(x) + (\lambda_0 + d(x))\,\psi(x) = \psi(0)b(x), \quad x \geq 0, \\ \psi(x) \geq 0, \quad \int_0^\infty \psi(x)\varphi(x)dx = 1. \end{cases}$$

然后，我们定义重整化密度 $\tilde{n}(t,x) := n(t,x)e^{-\lambda_0 t}$。证明，对于任意凸函数 $H : \mathbb{R}^+ \to \mathbb{R}^+$ 满足 $H(0) = 0$, 我们有

$$\frac{d}{dt}\int_0^\infty \psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)dx \leq 0, \quad \forall t \geq 0,$$

并证明

$$\int_0^\infty \psi(x)n(t,x)dx = e^{\lambda_0 t}\int_0^\infty \psi(x)n_0(x)dx.$$

为了简化证明，我们在演算中假定在 ∞ 处的边界项的贡献都是可以忽略的。

R8 答案: (i) 这个方程推导方式有很多。举两个例子。

1，特征线法。由于使用时长随时间线性增长，我们定义特征线 $x(t)$，它满足

$$\frac{dx(t)}{dt} = 1.$$

而顺着特征线，根据停止速率的含义，我们有

$$\frac{d}{dt}n(t, x(t)) = -d(x(t))n(t, x(t)).$$

整理即得(7)式方程。

2, 微元法。考虑一个时间微元 $\delta t \ll 1$，根据假设一和假设二，我们有

$$n(t + \delta t, x + \delta t) = n(t, x) - \delta t d(x)n(x, t) + o(\delta t).$$

其中右端第一项表示时间平移的贡献，第二项表示停止的客户数量。两边除以 δt，再令 $\delta t \to 0$，即得此方程。

关于 $N(t)$ 的定义，只需要说明老客户推荐的贡献。对于固定某个使用时长 x 的老客户，他们单位时间内介绍的新客户的数量为 $b(x)n(t, x)$。为了求单位时间内所有老客户介绍的新客户数量，需要把所有使用时长的老客户的贡献加在一起，故表达为 $\int_0^\infty b(y)n(t, y)dy$。

(ii) 根据题意和 $N(t)$ 的定义，我们需要先把密度函数 $n(t, x)$ 写成 $N(t)$ 和其他参数的表达式，这需要求解此方程。

注意到这是一个一阶双曲方程，可以用特征线法求解。将方程改写为

$$\frac{d}{ds}n(t + s, x + s) + d(x + s)n(t + s, x + s) = 0,$$

则如果定义 $D(x) = \int_0^x d(y)dy$，那么

$$\frac{d}{ds}\left[e^{D(x+s)}n(t + s, x + s)\right] = 0. \tag{9}$$

那么，当 $s \geq \max(-t, -x)$ 时，我们有

$$e^{D(x+s)}n(t + s, x + s) = e^{D(x)}n(t, x), \quad \forall x \geq 0, t \geq 0. \tag{10}$$

特别的，我们可以令 $x = y$，$s = -y$ 可得，当 $t \geq y$ 时

$$n(t, y) = N(t - y)e^{-D(y)}.$$

再令 $x=y$，$s=-t$ 可得，当 $t\le y$ 时

$$n(t,y)=n_0(y-t)e^{D(y-t)-D(y)}.$$

为了导出 $N(t)$ 满足的方程，我们将它的表达式拆分成两部分

$$N(t)=\int_0^\infty b(y)n(t,y)dy=\int_0^t b(y)n(t,y)dy+\int_t^\infty b(y)n(t,y)dy.$$

根据特征线可知，右端第一项的特征线起源于 $x=0,t\ge 0$，而第二项的特征线起源于 $x\ge 0,t=0$。将 $n(t,y)$ 的表达式分别代入，我们得到

$$N(t)=\int_0^t b(y)e^{-D(y)}N(t-y)dy+\int_t^\infty b(y)e^{D(y-t)-D(y)}n_0(y-t)dy. \tag{11}$$

整理，即得到 $N(t)$ 满足的方程

$$N(t)=\int_0^t b(t-x)e^{-D(t-x)}N(x)dx+\int_0^\infty b(x+t)e^{D(x)-D(x+t)}n_0(x)dx. \tag{12}$$

考虑到 $d(x)>0$ 所以 D 是递增函数，上式中的 $e^{-D(t-x)}$, $e^{D(x)-D(x+t)}$ 均不大于 1。于是,再利用 $b(x)$ 的有界性，我们可以对 $N(t)$ 做如下估计

$$|N(t)|\le \|b\|_\infty\int_0^t |N(x)|dx+\|b\|_\infty\int_0^\infty |n_0(x)|\,dx.$$

最后，利用Gronwall 引理，我们就可以得到待证的不等式。
(iii) 这是一个广义的相对熵估计的问题。
首先我们将方程(7)改写成重整化密度函数满足的方程

$$\frac{\partial}{\partial t}\widetilde{n}(t,x)+\frac{\partial}{\partial x}\widetilde{n}(t,x)+(\lambda_0+d(x))\,\widetilde{n}(t,x)=0.$$

然后，我们进一步地改写，整理得到

$$\frac{\partial}{\partial t}\frac{\tilde{n}(t,x)}{\varphi(x)}+\frac{\partial}{\partial x}\frac{\tilde{n}(t,x)}{\varphi(x)}=0,$$

进而有

$$\frac{\partial}{\partial t}H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)+\frac{\partial}{\partial x}H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)=0.$$

将特征值问题和其对偶问题整合在一起，我们有

$$
\begin{cases}
\frac{\partial}{\partial x}[\varphi(x)\psi(x)] = -\psi(0)b(x)\varphi(x), \quad x \geq 0, \\
\psi(x) \geq 0, \quad \int_0^\infty \psi(x)\varphi(x)dx = 1.
\end{cases}
$$

通过直接计算，我们得到

$$
\frac{\partial}{\partial t}\left[\psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)\right] + \frac{\partial}{\partial x}\left[\psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)\right] = -\psi(0)b(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right).
$$

记 $d\mu(x) = b(x)\varphi(x)dx$，将上式对 x 在 $\mathbb{R}^+$ 上积分，可得,

$$
\frac{d}{dt}\int_0^\infty \psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)dx = -\psi(0)\int_0^\infty H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)d\mu(x) + \psi(0)H\left(\frac{\tilde{n}(t,0)}{\varphi(0)}\right).
$$

我们注意到，由定义知 $\varphi(0) = 1$，且 $n(t,0) = \int_0^\infty b(x)n(t,x)dx$，那么

$$
\frac{\tilde{n}(t,0)}{\varphi(0)} = \tilde{n}(t,0) = \int_0^\infty b(x)\tilde{n}(t,x)dx = \int_0^\infty \frac{\tilde{n}(t,x)}{\varphi(x)}d\mu(x),
$$

于是可得

$$
\frac{d}{dt}\int_0^\infty \psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)dx = \psi(0)\left[-\int_0^\infty H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)d\mu(x) + H\left(\int_0^\infty \frac{\tilde{n}(t,x)}{\varphi(x)}d\mu(x)\right)\right].
$$

再由Jensen不等式，我们可得

$$
\frac{d}{dt}\int_0^\infty \psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)dx \leq 0.
$$

最后，令 $H(u) = u$，即得待证等式。

Alibaba Global Mathematics Competition-Quanlifying

1,2. In a fictional world, each resident (viewed as geometric point) is assigned a number: $1, 2, \cdots$. In order to fight against some epidemic, the residents take some vaccine and they stay at the vaccination site after taking the shot for observation. Now suppose that the shape of the Observation Room is a circle of radius $\frac{1}{4}$, and one requires that the distance $d_{m,n}$ between the Resident No. m and the Resident No. n must satisfy

$$(m+n)d_{m,n} \geq 1.$$

Where we consider the distance on the circle, i.e., the length of the minor arc between two points.

1. **Multiple-Choice Question (4 points)** Which of the following is correct?

   A No more than 8 residents can be placed inside the observation room;

   B The maximal number of residents that can be placed simultaneously is greater than 8, but still finite;

   C Any number of residents can be placed inside the observation room.

2. **Proof Question (6 points)** Give a proof of your answer to Question (i).

R1 Answer. The Choice C is correct.

R2 Answer. Solution I. We can place the Residents No. $1, 2, \ldots$ according to the following rule. First, put Resident No.1 arbitrarily. For $n > 2$, if Residents No.$1, 2, \ldots, n-1$ have already been placed, we consider the positions where Resident No. n cannot be placed.

For $1 \le m \le n-1$, by $d_{m,n} \ge \frac{1}{m+n}$, we know that the Resident No. n cannot be placed in the arc that is centered at Resident No. m, and of the length $\frac{2}{m+n}$. The total length of these arcs is

$$\frac{2}{n+1}+\frac{2}{n+2}+\cdots+\frac{2}{2n-1} < 2(\ln\frac{n+1}{n}+\ln\frac{n+2}{n+1}+\cdots+\ln\frac{2n-1}{2n-2}) = 2\ln\frac{2n-1}{n} < 2\ln 2.$$

Therefore, the total length of the union of these arcs does not exceed $2\ln 2$, while the perimeter of the circle is $\frac{1}{4}\cdot 2\pi = \frac{\pi}{2}$. It is easy to observe that $\frac{\pi}{2} > 1.5 > 2\ln 2$, so these arcs would not cover the whole circle, hence it is always possible to find a place for Resident No. n such that its distances to Residents No.$1, 2, \ldots, n-1$ satisfy the requirement. By induction we conclude that the circle can accomodate any quantity of residents.

Solution II. We consider the Cartesion coordinate system whose origin is the center of the circle, and place Residents No.$1, 2, 3$ and 4 at $(\frac{1}{4}, 0), (-\frac{1}{4}, 0), (0, \frac{1}{4}), (0, -\frac{1}{4})$, respectively, or in an equivalent way, we say that (the principle values of) their arguments are $0, \pi, \frac{\pi}{2}$ and $\frac{3\pi}{2}$. Now the distance between any two residents is no less than $2\cdot\frac{1}{4}\cdot\frac{\pi}{4} = \frac{\pi}{8} > \frac{1}{1+2}$, so the placement of these 4 residents satisfies the requirement.

We prove the following assertion by induction: for any integer $k > 2$, we can place Residents No. $1, 2, \ldots, 2^k$ at the vertices of a regular 2^k-gon inscribed to the circle, such that their mutual distances fulfill the requirement, and no two residents among the first 2^{k-1} occupy adjacent vertices.

The assertion holds for $k = 2$. If it is valid for k, so that the first 2^k residents are placed. They divide the circle into 2^k equal arcs. We need to put Residents No. $2^k+1, 2^k+2, \ldots, 2^{k+1}$ at the midpoints of these arcs. So we just need to prove the distances related to Residents No. $2^k+1, 2^k+2, \ldots, 2^{k+1}$ satisfy the requirement.

We put Residents No. $2^k+1, 2^k+2$ in positions next to Resident No. 2^{k-1} (i.e. the corresponding arguments differ to that of Resident No. 2^{k-1} by $\frac{2\pi}{2^{k+1}}$); put Residents No. $2^k+3, 2^k+4$ next to Resident No. $2^{k-1}-1$; $\cdots$ put Residents No. $2^k+2a-1, 2^k+2a$ next to Resident No. $2^{k-1}-a+1$; $\cdots$ put Residents No. $2^{k+1}-1, 2^{k+1}$ next to Resident No.1.

As the first 2^{k-1} residents do not occupy any consecutive positions, the above placement would not cause any problem. Now we consider distance between two residents (only the cases where at least one resident in the pair is "new" need to be considered).

Since the circle is now divided into 2^{k+1} arcs, each has length $\frac{2\pi}{2^{k+1}} \cdot \frac{1}{4} = \frac{\pi}{2^{k+2}}$. For Residents No. $m(> 2^k)$ and n, if they are separated by at least two pieces of arcs, then

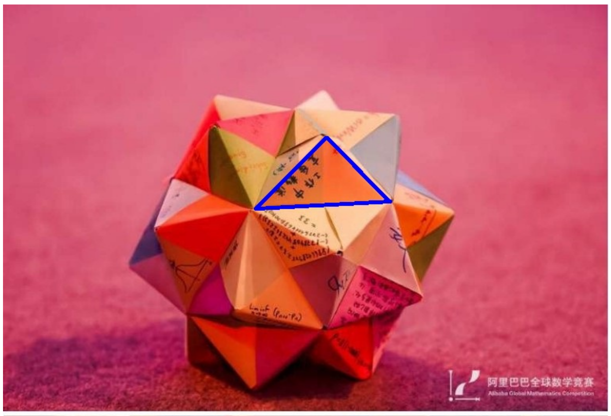
$$d_{m,n} \geq \frac{\pi}{2^{k+1}} > \frac{\pi}{2} \cdot \frac{1}{m+n} \cdot \frac{2(m+n)}{2^{k+1}} > \frac{1}{m+n};$$

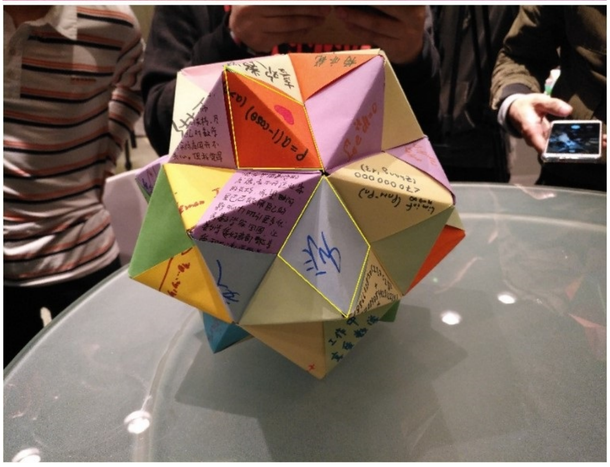
If they are just separated by one piece of arc, then for $n \in \{2^k + 2a - 1, 2^k + 2a\}$, we have $m > 2^{k-1} - a + 1$, hence

$$(m+n)d_{m,n} \geq \frac{\pi}{2^{k+2}} \cdot (2^k + 2a - 1 + 2^{k-1} - a + 1) = \frac{\pi}{2^{k+2}} \cdot (2 \cdot 2^{k-1} + a) \geq \frac{3\pi \cdot 2^{k-1}}{2^{k+1}} = \frac{3\pi}{8} > 1.$$

Therefore, the distances between each pair of Residents No. $1, 2, \ldots, 2^{k+1}$ satisfies the requirement. Then by induction, we conclude that any number of residents can be accomodated in that way.

3,4. Two years ago, the winners of 2018 Alibaba Global Mathematics Competition made a paper polyhedron as a present to the organizers. As shown in the photoes, the polyhedron has 60 equal triangular faces, and its surface can also be divided into 60 congruent non-planar quadrilaterals.





A non-planar n-gon is a non-planar figure obtained from a planar n-gon by folding it along some diagonals (i.e., to form certain dihedral angle at each of the chosen diagonals). Two non-planar figures are congruent if and only if each can be obtained from the other one by a certain isometry of $\mathbb{R}^3$. A polyhedron is a bounded region in $\mathbb{R}^3$ whose boundary is the union of a finite collection of planar polygons along common edges.

3. **True-False Question (4 points)** We know that $2021 = 43 \times 47$. Is there a polyhedron whose surface can be formed by gluing together 43 equal non-planar 47-gons?

4. **Question and Answer (6 points)** Please justify your answer to Question (i) with a rigorous argument.

R3,Answer. The answer is YES.

R4 Answer. All we need to do is to construct an example. Let's consider a standard torus $\mathbb{T}$, whose points can be represented by two parameters:

$$\mathbb{T} = \{\theta, \varphi : \ 0 \leq \theta, \varphi < 2\pi\}.$$

One can view the z-axis as the axis of symmetry of the torus:

$$((R + r\cos\varphi)\cos\theta, (R + r\cos\varphi)\sin\theta, r\sin\varphi).$$

For $1 \leq k \leq 43$, we consider the following region on the torus

$$D_k = \{\theta, \varphi : \ \frac{2(k-1)}{43}\pi + 3\frac{\varphi}{86} \leq \theta \leq \frac{2k}{43}\pi + 3\frac{\varphi}{86}\}.$$

Intuitively, what we do here is to divide the torus into 43 equal parts, then cut every part along the circle $\{\varphi = 0\}$, keep one side of the cut while sliding the other side along the circle for certain angle.

Now, we deform the circle $\{\varphi = 0\}$ into a regular 43-gon whose vertices correspond to $\theta = \frac{2k}{43}\pi$. Then D_k has four "sides" of (two of which lie on $\{\varphi = 0\}$), four "corners" (two of which are adjacent vertices of the 43-gon, while the other two are midpoints of two sides, we need then mark the vertex of the 43-gon between these two midpoints). We denote

$$\begin{aligned}
C_{k,0} &= (\frac{2(k-1)}{43}\pi, 0), C_{k,1} = (\frac{2k}{43}\pi, 0);\\
D_{k,0} &= (\frac{2k+1}{43}\pi, 2\pi), D_{k,1} = (\frac{2k+3}{43}\pi, 2\pi);\\
E_k &= (\frac{2k+2}{43}\pi, 2\pi).
\end{aligned}$$

Take another "side" of ∂D_k, mark 21 points, e.g.

$$A_{k,i} = (\frac{2(k-1)}{43}\pi + 3\frac{\varphi}{86}\pi, \frac{i}{11}\pi), i = 1, \ldots, 21.$$

Then rotate around z-axis by $\frac{2}{43}\pi$ to get another 21 points, denote them by $B_{k,i}, i = 1, \ldots, 21$.

Now we join

$$C_{k,0}C_{k,1}, C_{k,0}A_{k,1}, C_{k,1}B_{k,1}, A_{k,i}A_{k,i+1}, B_{k,i}B_{k,i+1}, A_{k,i}B_{k,i}, A_{k,i}B_{k,i+1} \ (i = 1, \ldots, 21),$$

and $A_{k,21}D_{k,0}, B_{k,21}D_{k,1}, A_{k,21}E_k, B_{k,21}E_k, D_{k,0}E_k, E_kD_{k,1}$. We get a non-planar 47-gon. Thus we get 43 congruent (the construction above is independent of k)non-planar 47-gons, they can be glue together to form a polyhedron.

Remark: A typical mistake would be to think that the vertices(edges) of these non-planar polygons are all vertices(edges, not part of eadges) of the polyhedron, then d-educts from "each edge is counted twice" and "2021 is an odd number" a fake "condtra-diction".

5. Last year, Master Zhang got famous with his topology noodle, and he is now a part-time chef at Alibaba Damo Academy. One day, Xiao Li, a software engineer at the DAMO Academy, complained about his recent work with Master Zhang. Xiao Li's job involves designing algorithms to tune the parameters of various systems, which can usually be formulated as finding a vector $\mathbf{x} \in \mathbb{R}^n$ that minimizes a certain loss function on $\mathbb{R}^n$. In his latest project, Xiao Li has to deal with a loss function f that is provided by another group. For security and technical reasons, the other group cannot reveal the explicit definition f, but only offers an interface to evaluate f at any given $\mathbf{x} \in \mathbb{R}^n$. Hence, Xiao Li needs to minimize f based on its function values. In addition, the evaluation of f is costly. Fortunately, the dimension n of this problem is not high (around 10). Besides, the group that provides the function informs Xiao Li that he may assume f is smooth.

Xiao Li's problem reminds Master Zhang about an antique radio that he owns. To listen to a program on this radio, you need to carefully tune the frequency knob forward and backward while monitoring the quality of the sound until finding the best frequency for receiving the signal. In this process, nobody knows the precise relation between the angle of the knob and the quality of the sound. After a careful discussion with Master Zhang, Xiao Li realizes that minimizing f is like tuning a machine with multiple knobs: just imagine that each component of $\mathbf{x}$ is controlled by a knob, and $f(\mathbf{x})$ represents certain performance of the machine; one should be able to find the best $\mathbf{x}$ by tuning each knob forward and backward while monitoring the value of f. Therefore, Xiao Li and Master Zhang propose an iterative algorithm for minimizing f, named Automated Forward/Backward Tuning (AFBT, Algorithm 1). At iteration k, AFBT considers $2n$ points $\{\mathbf{x}_k \pm t_k\mathbf{e}^i : i = 1, \ldots, n\}$ by varying each component of $\mathbf{x}_k$ forward or backward with a step size t_k, sets $\mathbf{y}_k$ to the one rendering the smallest value of f, and checks whether $\mathbf{y}_k$ achieves a sufficient decrease in f; if yes, it takes $\mathbf{x}_{k+1} = \mathbf{y}_k$ and doubles the step size; otherwise, it sets $\mathbf{x}_{k+1} = \mathbf{x}_k$ and halves the step size. In the algorithm, $\mathbf{e}^i$ denotes the i-th canonical coordinate vector in $\mathbb{R}^n$ (the i-th entry is 1 while all the others are 0); $\mathbb{1}(\cdot)$ is the indicator function, so that $\mathbb{1}[f(\mathbf{x}_k) - f(\mathbf{y}_k) \geq t_k^2]$ equals 1 if $f(\mathbf{x}_k) - f(\mathbf{y}_k)$ is at least the square of t_k, or else the value is 0.

Algorithm 1 Automated Forward/Backward Tuning (AFBT)

Input $\mathbf{x}_0 \in \mathbb{R}^n$ and $t_0 > 0$. For $k = 0, 1, 2, \ldots$, do the following.

1: $\mathbf{y}_k := \operatorname{argmin}\left\{f(\mathbf{y}) : \mathbf{y} = \mathbf{x}_k \pm t_k\mathbf{e}^i,\ i = 1, \ldots, n\right\}$ `# Evaluate loss function.`

2: $s_k := \mathbb{1}\left[f(\mathbf{x}_k) - f(\mathbf{y}_k) \geq t_k^2\right]$ `# Sufficient decrease? Yes:` $s_k = 1$; `No:` $s_k = 0$.

3: $\mathbf{x}_{k+1} := (1 - s_k)\mathbf{x}_k + s_k\mathbf{y}_k$ `# Update iterate.`

4: $t_{k+1} := 2^{2s_k - 1}t_k$ `# Update step size.` $s_k = 1$: `double;` $s_k = 0$: `halve.`

Now we make the following assumptions on the loss function $f : \mathbb{R}^n \to \mathbb{R}$.

Assumption 1. f is convex. This means that

$$f((1-\alpha)\mathbf{x}+\alpha\mathbf{y}) \;\le\; (1-\alpha)f(\mathbf{x})+\alpha f(\mathbf{y})$$

for all $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ and $\alpha\in[0,1]$.

Assumption 2. f is differentiable on $\mathbb{R}^n$ and ∇f is L-Lipschitz on $\mathbb{R}^n$.

Assumption 3. f is level-bounded, meaning that $\{\mathbf{x}\in\mathbb{R}^n \mid f(\mathbf{x})\le\lambda\}$ is a bounded set for any given $\lambda\in\mathbb{R}$.

Based on **Assumptions 1** and **2**, we can prove that

$$\langle\nabla f(\mathbf{x}),\ \mathbf{y}-\mathbf{x}\rangle \;\le\; f(\mathbf{y})-f(\mathbf{x}) \;\le\; \langle\nabla f(\mathbf{x}),\ \mathbf{y}-\mathbf{x}\rangle+\frac{L}{2}\|\mathbf{x}-\mathbf{y}\|^2$$

for all $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$; **Assumptions 1** and **3** ensure that f has a finite minimum f^* on $\mathbb{R}^n$. For more properties of convex functions, see any textbook on convex analysis.

Proof Question (20 points) Under **Assumptions 1–3**, prove for **AFBT** that

$$\lim_{k\to\infty} f(\mathbf{x}_k) \;=\; f^*.$$

R5 Proof. Assume that $f(\mathbf{x}_k) \nrightarrow f^*$. Then $\inf_{k\geq 0}[f(\mathbf{x}_k) - f^*] > 0$ since $\{f(\mathbf{x}_k)\}$ is non-increasing, Denote $\mathbf{g}_k = \nabla f(\mathbf{x}_k)$. Then $\inf_{k\geq 0} \|\mathbf{g}_k\| > 0$ (Pick an $\mathbf{x}^*$ with $f(\mathbf{x}^*) = f^*$. Then $f(\mathbf{x}_k) - f^* \leq \langle \mathbf{g}_k, \mathbf{x}_k - \mathbf{x}^* \rangle$ by the convexity of f; meanwhile, $\{\mathbf{x}_k\}$ is bounded due to the monotonicity of $\{f(\mathbf{x}_k)\}$ and the level-boundedness of f. Thus $\inf_{k\geq 0} \|\mathbf{g}_k\| > 0$). In other words, there exists an $\varepsilon > 0$ such that $\|\mathbf{g}_k\| \geq \varepsilon$ for all $k \geq 0$. Given any $k \geq 0$, we can pick an $i_k \in \{1, \ldots, n\}$ satisfying

$$|\langle \mathbf{g}_k, \mathbf{e}^{i_k} \rangle| \;\geq\; \kappa \|\mathbf{g}_k\| \;\geq\; \kappa\varepsilon \tag{1}$$

with $\kappa = 1/\sqrt{n}$. Hence

$$f(\mathbf{y}_k) \;\leq\; \min\{f(\mathbf{x}_k \pm t_k \mathbf{e}^{i_k})\} \;\leq\; f(\mathbf{x}_k) - t_k |\langle \mathbf{g}_k, \mathbf{e}^{i_k} \rangle| + \frac{L t_k^2}{2} \;\leq\; f(\mathbf{x}_k) - \kappa\varepsilon t_k + \frac{L t_k^2}{2}. \tag{2}$$

Therefore, whenever

$$t_k \;\leq\; \frac{2\kappa\varepsilon}{L+2}, \tag{3}$$

we will have $f(\mathbf{y}_k) \leq f(\mathbf{x}_k) - t_k^2$, which renders $s_k = 1$ and hence $t_{k+1} = 2t_k$. It is then easy to see that $\{t_k\}$ has a positive lower bound, namely

$$t_k \;\geq\; \underline{t} \;\equiv\; \min\left\{ t_0,\; \frac{\kappa\varepsilon}{L+2} \right\} \;>\; 0 \qquad \text{for each} \qquad k \geq 0. \tag{4}$$

Thus $s_k = 1$ for infinitely many k (otherwise, $t_k \to 0$), and $f(\mathbf{x}_k) - f(\mathbf{x}_{k+1}) \geq \underline{t}^2$ for each of such k, contradicting the lower-boundedness of f. The proof is complete.

6. Let n be a positive integer. For any positive integer k, write $0_k = \mathrm{diag}\{\underbrace{0,\ldots,0}_{k}\}$ for the $k \times k$ zero matrix. Let

$$Y = \begin{pmatrix} 0_n & A \\ A^t & 0_{n+1} \end{pmatrix}$$

be a $(2n+1)\times(2n+1)$ matrix, where $A = (x_{i,j})_{1\le i\le n, 1\le j\le n+1}$ is an $n\times(n+1)$ real matrix and A^t denotes the transpose of A, that is, the $(n+1)\times n$ matrix whose (j,i)-entry is $x_{i,j}$.

(i) **Proof Question (10 points)** A complex number λ is called an eigenvalue of a $k \times k$ matrix X if $Xv = \lambda v$ for some nonzero column vector $v = (x_1, \ldots, x_k)^t$. Show that: 0 is an eigenvalue of Y and every other eigenvalue of Y is of the form $\pm\sqrt{\lambda}$ where λ is a non-negative real eigenvalue of AA^t.

(ii) **Proof Question (15 points)** Let $n = 3$ and a_1, a_2, a_3, a_4 be four distinct positive real numbers. Put

$$a = \sqrt{\sum_{1\le i\le 4} a_i^2}$$

and

$$x_{i,j} = a_i\delta_{i,j} + a_j\delta_{4,j} - \frac{1}{a^2}(a_i^2 + a_4^2)a_j$$

$(1 \le i \le 3,\ 1 \le j \le 4)$, where $\delta_{i,j} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \ne j \end{cases}$. Show that: Y has 7 distinct eigenvalues.

R6 Proof. **(i)** Write $I_n = \mathrm{diag}\{\underbrace{1, \ldots, 1}_{n}\}$ for the $n \times n$ identity matrix. By an elementary reduction the characteristic polynomial

$$\det(\lambda I_{2n+1} - Y) = \lambda \det(\lambda^2 I_n - AA^t).$$

Then, 0 is an eigenvalue of Y and every other eigenvalue of Y is of the form $\pm\sqrt{\lambda}$ where λ is a non-negative real eigenvalue of AA^t.
(ii) Put $u = (a_4, a_4, a_4)$ and $v = (\frac{a_1^2+a_4^2}{a}, \frac{a_2^2+a_4^2}{a}, \frac{a_3^2+a_4^2}{a})$. By calculation we have

$$AA^t = \mathrm{diag}\{a_1^2, \ldots, a_3^2\} + u^t u - v^t v.$$

Let $f(s) = \det(sI_n - AA^t)$ be the characteristic polynomial of AA^t. Then, by calculation one shows that

$$f(a_i^2) = \frac{a_i^2}{a^2} \prod_{1 \le j \le 4, j \ne i} (a_i^2 - a_j^2)$$

for each $i \in \{1, 2, 3, 4\}$. Let a_1', a_2', a_3', a_4' be the descending re-ordering of a_1, a_2, a_3, a_4. Then, we get: AA^t has three distinct eigenvalues b_1^2, b_2^2, b_3^3 where b_1, b_2, b_3 are positive real numbers such that

$$a_1' > b_1 > a_2' > b_2 > a_3' > b_3 > a_4'.$$

Then, by (i) Y has 7 distinct real eigenvalues.

7. For a continuous and absolutely integrable complex-valued function $f(x)$ on $\mathbb{R}$, define a function $(Sf)(x)$ on $\mathbb{R}$ by

$$(Sf)(x) = \int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} ux} f(u) \mathrm{d}u.$$

(i) **Question and Answer (10 points)** Find explicit forms of $S(\frac{1}{1+x^2})$ and $S(\frac{1}{(1+x^2)^2})$.

(ii) **Question and Answer (15 points)** For any integer k, write $f_k(x) = (1+x^2)^{-1-k}$. When $k \geq 1$, find constants c_1, c_2 such that the function $y = (Sf_k)(x)$ solves a second order differential equation

$$xy'' + c_1 y' + c_2 xy = 0.$$

R7 Answer. Write V for the space of complex-valued, continuous and absolutely integrable functions on $\mathbb{R}$.

Lemma 0.1. *(i) If $f(x) \in V$, $f'(x) \in V$ and $\lim_{x\to\infty} f(x) = 0$, then*

$$(Sf')(x) = -2\pi \mathbf{i} x (Sf)(x). \tag{5}$$

(ii) If $f(x) \in V$ and $xf(x) \in V$, then

$$(Sf)' = 2\pi \mathbf{i} S(xf(x)). \tag{6}$$

Proof. (i)

$$\begin{aligned}
& (Sf')(x) \\
= \; & \int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} ux} f'(u) \mathrm{d}u \\
= \; & e^{2\pi \mathbf{i} ux} f(u)|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} (e^{2\pi \mathbf{i} ux})' f(u) \mathrm{d}u \\
= \; & -2\pi \mathbf{i} x \int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} ux} f(u) \mathrm{d}u \\
= \; & -2\pi \mathbf{i} x (Sf)(x)
\end{aligned}$$

(ii) For any $a, b \in \mathbb{R}$ with $a < b$,

$$\begin{aligned}
& \int_a^b 2\pi \mathbf{i} S(xf(x)) \mathrm{d}x \\
= \; & \int_a^b 2\pi \mathbf{i} (\int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} ux} u f(u) \mathrm{d}u) \mathrm{d}x \\
= \; & \int_{-\infty}^{+\infty} (\int_a^b 2\pi \mathbf{i} u e^{2\pi \mathbf{i} ux} f(u) \mathrm{d}x) \mathrm{d}u \\
= \; & \int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} bu} f(u) \mathrm{d}u - \int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} au} f(u) \mathrm{d}u \\
= \; & (Sf)(b) - (Sf)(a).
\end{aligned}$$

Thus, $(Sf)' = 2\pi \mathbf{i} S(xf(x))$. □

Lemma 0.1 implies the following corollary immediately.

Corollary 0.2. *(i) Assume that $f, f', Sf, x(Sf)(x) \in V$ and*

$$\lim_{x\to\infty} f(x) = 0.$$

If $(S(Sf))(x) = f(-x)$, then $(S(Sf'))(x) = f'(-x)$.

(ii) Assume that $f(x), xf(x), Sf, (Sf)' \in V$ *and*

$$\lim_{x\to\infty}(Sf)(x) = 0.$$

If $(S(Sf))(x) = f(-x)$*, then* $S(S(xf(x))) = -xf(-x)$*.*

Lemma 0.3. *(i)* $S((1+x^2)^{-1}) = \pi e^{-2\pi|x|}$*.*

(ii) $S(\pi e^{-2\pi|x|}) = (1+x^2)^{-1}$*.*

Proof. (i) Write $f(x) = (1+x^2)^{-1}$. For $x \geq 0$, we have

$$(Sf)(x) = \lim_{A\to+\infty}\int_{-A}^{A}\frac{e^{2\pi \mathbf{i} ux}}{1+u^2}\mathrm{d}u.$$

Put

$$C_A := \{z = u + \mathbf{i}v : -A \leq u \leq A, v = 0\}\bigcup\{z = Ae^{\mathbf{i}\theta} : 0 \leq \theta \leq \pi\}.$$

Note that, $\mathbf{i}$ is the only pole of $\frac{1}{1+z^2}$ inside the domain bounded by C_A whenever $A > 1$. Using the trick of contour integral and letting $A \to \infty$, we get $(Sf)(x) = \pi e^{-2\pi x}$. Since $f(x)$ is an even function, so is $(Sf)(x)$. Then, $(Sf)(x) = \pi e^{-2\pi|x|}$.

(ii) Write $g(x) = \pi e^{-2\pi|x|}$. By direct calculation

$$\begin{aligned}
& (Sg)(x) \\
= & \int_{-\infty}^{\infty} e^{2\pi \mathbf{i} xu}\pi e^{-2\pi|u|}\mathrm{d}u \\
= & \pi\int_{0}^{\infty}(e^{2\pi \mathbf{i} xu} + e^{-2\pi \mathbf{i} xu})e^{-2\pi u}\mathrm{d}u \\
= & -\frac{1}{2}(\frac{e^{-2\pi(1+\mathbf{i}x)u}}{1+\mathbf{i}x} + \frac{e^{-2\pi(1-\mathbf{i}x)u}}{1-\mathbf{i}x})|_0^\infty \\
= & \frac{1}{1+x^2}.
\end{aligned}$$

□

Lemma 0.4. *(i) For any* $k \geq 0$*,* Sf_k *is of the form* $(Sf_k)x = e^{-2\pi|x|}g_k(|x|)$ *where* g_k *is a polynomial of degree* k*.*

(ii) For any $k \geq 0$*,* $S(S(f_k)) = f_k$ *and* $S(S(xf_{k+1}(x))) = -xf_{k+1}(x)$*.*

Proof. (i) We have a recursive relation

$$f_{k+1} = f_k + \frac{1}{2(k+1)}xf_k'(x).$$

By Lemma 0.1, we get a recursive relation for Sf_k:

$$Sf_{k+1} = Sf_k - \frac{1}{2(k+1)}(x(Sf_k)(x))'.$$

Then, one proves the conclusion by induction.

(ii) Note that $f_k'(x) = -2(k+1)xf_{k+1}(x)$ and $(xf_k(x))' = -(1+2k)f_k(x)+2(k+1)f_{k+1}(x)$. By the conclusion of part (i) and Lemma 0.1, we see that the assumptions in Corollary 0.2 are all satisfied for $f_k(x)$ and $xf_{k+1}(x)$ $(k \geq 0)$. Then, one shows by induction that $S(S(f_k)) = f_k(x)$ and $S(S(xf_{k+1}(x))) = -xf_{k+1}(x)$ for any $k \geq 0$. $\square$

Back to the problem. (i) It is shown in Lemma 0.3 that $S((1+x^2)^{-1}) = \pi e^{-2\pi|x|}$. By (5), we get

$$S(-2x(1+x^2)^{-2}) = -2\pi^2 \mathbf{i} x e^{-2\pi|x|}.$$

By (6), we get

$$S(-2x^2(1+x^2)^{-2}) = -\pi(1-2\pi|x|)e^{-2\pi|x|}.$$

Then,

$$S((1+x^2)^{-2}) = \frac{\pi}{2}(1+2\pi|x|)e^{-2\pi|x|}.$$

(ii) Firstly, $x^j f_k(x)$ $(0 \leq j \leq 2k)$ are all absolutely integrable when $k \geq 1$. Then, by (6), $y = (Sf_k)(x)$ is a $2k$-th order continuous differentiable function. By Lemma 0.1 and Lemma 0.4, $xy'' + c_1 y' + c_2 xy = 0$ is equivalent to

$$(x^2 f_k' + 2xf_k) - c_1 x f_k - \frac{c_2}{4\pi^2} f_k' = 0.$$

Inputting $f_k(x) = (1+x^2)^{-1-k}$, we get $c_1 = -2k$ and $c_2 = -4\pi^2$.

8. When a new social networking software appears on the market, the software company cares about how the number of the active customers increases in time, and also would like to investigate how certain traits of the customers evolve over time. We denote the population density of the customers by $n(t, x)$, where t is the time variable and x represents how long an active customer has been using this software . We assume that the time evolution of the density $n(t, x)$ is governed by the following factors:

Assumption 1. When a customer keeps using the software, his or her usage time length x increases linearly in time.

Assumption 2. When a customer uses the software, he or she may stop using it with a stopping rate $d(x) > 0$. Here, we assume the stopping rate only depends on x.

Assumption 3. There are two sources of new customers:

1. The software company promotes this software by commercial advertisements. The rate of change in the number of customers due to the advertisements is denoted by $c(t)$.
2. The existing customers may recommend this software to their family or friend-s. The effective recommendation rate is denoted by $b(x)$, which is related to the customer's usage time length.

We assume at $t = 0$, the population density is given, $n(0, x) = n_0(x)$. We can derive that the time evolution of $n(t, x)$ is given by

$$\begin{cases} \frac{\partial}{\partial t} n(t,x) + \frac{\partial}{\partial x} n(t,x) + d(x) n(t,x) = 0, \quad t \geq 0, x \geq 0, \\ N(t) := n(t, x = 0) = c(t) + \int_0^\infty b(y) n(t,y) dy. \end{cases} \tag{7}$$

Here, $N(t)$ is interpreted as the increasing rate of the new customers. We assume $b, d \in L^\infty_+(0, \infty)$, that is, $b(x)$ and $d(x)$ are positive and (essentially) bounded. From this point on, we also assume $c(t) \equiv 0$ for simplicity of analysis.

(i) **Question and Answer (10 points)** With **Assumption 1** and **Assumption 2**, formally derive the partial differential equation that $n(t, x)$ satisfies as in (7), and clearly indicate the correspondence between the assumptions and the mathematical expressions during the derivation. Also, explain the meaning of $N(t)$ as given in (7).

(ii) **Question and Answer (10 points)** We aim to establish the connection between $N(t)$ and $b(x)$. To fulfil this task, derive an equation that $N(t)$ satisfies, such that the equation only contains $N(t)$, $n_0(x)$, $b(x)$ and $d(x)$, but $n(t, x)$ does not appear in this equation. Prove that $N(t)$ satisfies the following estimate

$$|N(t)| \leq \|b\|_\infty e^{\|b\|_\infty t} \int_0^\infty |n_0(x)|\, dx, \tag{8}$$

where $\|\cdot\|_\infty$ denotes the L^∞ norm.

(iii) **Proof Question (10 points)** Finally, we aim to explore the long time asymptotic behavior of the population density $n(t,x)$. Since the total number might be increasing, it is more convenient to work with a normalized density function.

We further assume the following eigenvalue problem has a unique solution $(\lambda_0, \varphi(x))$

$$\begin{cases} \varphi'(x) + (\lambda_0 + d(x))\,\varphi(x) = 0, \quad x \geq 0, \\ \varphi(x) > 0, \quad \varphi(0) = \int_0^\infty b(x)\varphi(x)dx = 1, \end{cases}$$

and its dual problem has a unique solution $\psi(x)$:

$$\begin{cases} -\psi'(x) + (\lambda_0 + d(x))\,\psi(x) = \psi(0)b(x), \quad x \geq 0, \\ \psi(x) \geq 0, \quad \int_0^\infty \psi(x)\varphi(x)dx = 1. \end{cases}$$

We define the normalized density $\tilde{n}(t,x) := n(t,x)e^{-\lambda_0 t}$. Prove that for any convex $H : \mathbb{R}^+ \to \mathbb{R}^+$ with $H(0) = 0$, we have

$$\frac{d}{dt}\int_0^\infty \psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)dx \leq 0, \quad \forall t \geq 0,$$

and show that

$$\int_0^\infty \psi(x)n(t,x)dx = e^{\lambda_0 t}\int_0^\infty \psi(x)n_0(x)dx.$$

For simplicity, we assume that the related boundary terms at ∞ can be neglected.

R8 Answer. (i). There are many ways to derive the equation.

Approach 1: characteristic method. Since the usage time length grows linearly in time, we can define the characteristic $x(t)$ satisfying

$$\frac{dx(t)}{dt} = 1.$$

Along the characteristic and according to the meaning of the stopping rate, we have

$$\frac{d}{dt}n(t, x(t)) = -d(x(t))n(t, x(t)),$$

which leads to the PDE.

Approach 2: infinitesimal method. For an infinitesimal time $\delta t \ll 1$, by Assumption I and Assumption II, we have

$$n(t + \delta t, x + \delta t) = n(t, x) - \delta t d(x)n(x, t) + o(\delta t).$$

where the first term on the right hand side is from the time shifting, and the second term accounts for the stopping rate. If we divide each side by δt and let $\delta t \to 0$, then the PDE is derived.

For $N(t)$, it suffices to explain the contribution from the existing customers. For customers with a certain usage time length, their contribution to the increase rate of the new customers is $b(x)n(t, x)$. To obtain the increase rate from all existing customers, we should add up all their contributions, so the expression is given by $\int_0^\infty b(y)n(t, y)dy$.

(ii). We need to first solve the PDE to get an equation for $N(t)$.

This is a first order hyperbolic equation, which can be solved by the method of characteristics. We rewrite the equation as

$$\frac{d}{ds}n(t + s, x + s) + d(x + s)n(t + s, x + s) = 0,$$

then if we define $D(x) = \int_0^x d(y)dy$, we get

$$\frac{d}{ds}\left[e^{D(x+s)}n(t + s, x + s)\right] = 0. \tag{9}$$

When $s \geq \max(-t, -x)$, we have

$$e^{D(x+s)}n(t + s, x + s) = e^{D(x)}n(t, x), \quad \forall x \geq 0, t \geq 0. \tag{10}$$

In particular, we can let $x = y$ and $s = -y$ to get when $t \geq y$,

$$n(t, y) = N(t - y)e^{-D(y)}.$$

Let $x = y$, $s = -t$, we get when $t \leq y$,

$$n(t, y) = n_0(y - t)e^{D(y-t)-D(y)}.$$

In order to derive an equation for $N(t)$, we split its expression into two parts

$$N(t) = \int_0^\infty b(y)n(t,y)dy = \int_0^t b(y)n(t,y)dy + \int_t^\infty b(y)n(t,y)dy.$$

Clearly, the characteristics of the first term originate from $x = 0, t \geq 0$, and those of the second term originate from $x \geq 0, t = 0$. With the expression for $n(t,x)$, we get

$$N(t) = \int_0^t b(y)e^{-D(y)}N(t-y)dy + \int_t^\infty b(y)e^{D(y-t)-D(y)}n_0(y-t)dy, \tag{11}$$

which leads to

$$N(t) = \int_0^t b(t-x)e^{-D(t-x)}N(x)dx + \int_0^\infty b(x+t)e^{D(x)-D(x+t)}n_0(x)dx. \tag{12}$$

Since $d(x) > 0$, so D is an increasing function, and $e^{-D(t-x)}$, $e^{D(x)-D(x+t)}$ in the expression above are less than 1. By the boundedness of $b(x)$, we get

$$|N(t)| \leq \|b\|_\infty \int_0^t |N(x)|dx + \|b\|_\infty \int_0^\infty |n_0(x)|\, dx.$$

And with Gronwall's lemma, we can obtain the desired inequality.
(iii). This is a generalized relative entropy estimate.

First, we rewrite the original equation as an equation that the normalized density satisfies

$$\frac{\partial}{\partial t}\tilde{n}(t,x) + \frac{\partial}{\partial x}\tilde{n}(t,x) + (\lambda_0 + d(x))\,\tilde{n}(t,x) = 0.$$

We further rewrite

$$\frac{\partial}{\partial t}\frac{\tilde{n}(t,x)}{\varphi(x)} + \frac{\partial}{\partial x}\frac{\tilde{n}(t,x)}{\varphi(x)} = 0,$$

and then we get

$$\frac{\partial}{\partial t}H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right) + \frac{\partial}{\partial x}H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right) = 0.$$

With the eigenvalue problem and the dual problem, we get

$$\begin{cases} \frac{\partial}{\partial x}[\varphi(x)\psi(x)] = -\psi(0)b(x)\varphi(x), \quad x \geq 0, \\ \psi(x) \geq 0, \quad \int_0^\infty \psi(x)\varphi(x)dx = 1. \end{cases}$$

By direct calculations, we get

$$\frac{\partial}{\partial t}\left[\psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)\right] + \frac{\partial}{\partial x}\left[\psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)\right] = -\psi(0)b(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right).$$

Let $d\mu(x) = b(x)\varphi(x)dx$, and integrate the equation above over $\mathbb{R}^+$ w.r.t. x, we get

$$\frac{d}{dt}\int_0^\infty \psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)dx = -\psi(0)\int_0^\infty H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)d\mu(x) + \psi(0)H\left(\frac{\tilde{n}(t,0)}{\varphi(0)}\right).$$

We notice that $\varphi(0) = 1$, $n(t,0) = \int_0^\infty b(x)n(t,x)dx$, and thus

$$\frac{\tilde{n}(t,0)}{\varphi(0)} = \tilde{n}(t,0) = \int_0^\infty b(x)\tilde{n}(t,x)dx = \int_0^\infty \frac{\tilde{n}(t,x)}{\varphi(x)}d\mu(x),$$

then we get

$$\frac{d}{dt}\int_0^\infty \psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)dx = \psi(0)\left[-\int_0^\infty H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)d\mu(x) + H\left(\int_0^\infty \frac{\tilde{n}(t,x)}{\varphi(x)}d\mu(x)\right)\right].$$

By Jensen's inequality, we get

$$\frac{d}{dt}\int_0^\infty \psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)dx \le 0.$$

Finally by setting $H(u) = u$, we get the desired identity.