Generalization of Mrs. Gerber's lemma*

FAN CHENG

Mrs. Gerber's Lemma (MGL) hinges on the convexity of $H(p * H^{-1}(u))$, where $H(u)$ is the binary entropy function. In this work, we prove that $H(p*f(u))$ is convex in u for every $p \in [0,1]$ provided $H(f(u))$ is convex in u, where $f(u) : (a,b) \to [0, \frac{1}{2}]$. Moreover, our result subsumes MGL and simplifies the original proof. We show that the generalized MGL can be applied in binary broadcast channel to simplify some discussion.

KEYWORDS AND PHRASES: Mrs. Gerber's lemma, binary channel.

1. Introduction

Mrs. Gerber's Lemma (MGL) was introduced by Wyner and Ziv [1] in 1973, which was shown to be a binary version of the Entropy Power Inequality (EPI) by Shamai and Wyner [2]. In Witsenhausen [3], MGL was generalized to arbitrary binary input-output channels. In Ahlswede and Korner [4], they introduced the concept of the *gerbator* for arbitrary discrete memoryless channel to study MGL in alphabets with higher cardinality. In Chayat and Shamai [5], MGL was extended to arbitrary memoryless symmetric channels with binary inputs and discrete or continuous outputs. In Jog and Anantharam [6], they conjectured a strengthed MGL on an arbitrary abelian group and partially proved it. MGL is an instrumental tool to tackle the problems related to binary channels; e.g., the capacity region of binary symmetric broadcast channel (BS-BC) in Wyner [7]; the capacity region of BSC-BEC broadcast channel in Nair [8].

The rest of this work is organized as follows. In Section 2, we introduce the necessary notation and the background. In Section 3, we present our main result on the generalized MGL. In Section 4, we demonstrate the power of our result by simplifying the discussion in the binary broadcast channel.

2. Mrs. Gerber's lemma

For $x \in [0, 1]$, the binary entropy function is defined as

$$H(x) := -x \log x - (1 - x) \log(1 - x)^1$$

and the inverse of $H(x)$ is defined as

$$H^{-1}(x) \in [0, \frac{1}{2}].$$

Then

$$\frac{dH}{dx} = \log \frac{1 - x}{x} \text{ and } \frac{d^2 H}{dx^2} = -\frac{1}{x(1 - x)}.$$

The convolution of p and x is denoted by

$$p * x := p(1 - x) + (1 - p)x,$$

where $p \in [0, 1]$.

Theorem 1 (Mrs. Gerber's Lemma). *Let X be a Bernoulli random variable and let U be an arbitrary random variable. If $Z \sim \text{Bern}(p)$ is independent of (X, U) and $Y = X + Z \pmod 2$, then*

$$H^{-1}(H(Y|U)) \geq H^{-1}(H(X|U)) * p.$$

MGL can be equivalently proved via the following convexity lemma about the binary entropy function.

Lemma 1. *$H(p * H^{-1}(u))$ is convex in $u \in [0, 1]$ for every $p \in [0, 1]$.*

3. Generalization of MGL

We prove the following generalization of Mrs. Gerber's Lemma.

Theorem 2. *Let $f(u) : (a, b) \to [0, \frac{1}{2}]$ be twice differentiable. Then for every $p \in [0, 1]$, the function $H(p * f(u))$ is convex in u provided $H(f(u))$ is convex in u.*

---

[1] All logarithms in this work are natural.

Proof. The function $H(p * f(u))$ is symmetric in p about $\frac{1}{2}$, hence we can assume that $p \in [0, \frac{1}{2}]$. Since $f(u) \leq \frac{1}{2}$, $p * f(u) = (1 - 2p)f(u) + p \leq \frac{1}{2}$.

The second derivative of the given expression with respect to u is given by

$$(1) \qquad -\frac{((1 - 2p)f'(u))^2}{(1 - p * f(u))(p * f(u))} + (1 - 2p)f''(u) \log \frac{1 - p * f(u)}{p * f(u)}.$$

The convexity of $H(f(u))$ ($p = 0$ in (1)) implies that $f''(u) \geq 0$.

To show the convexity it suffices to show that

$$g(p) := -(1 - 2p)f'(u)^2 + (1 - p * f(u))(p * f(u))f''(u) \log \frac{1 - p * f(u)}{p * f(u)} \geq 0.$$

Further we know that at both $p = 0$ and $p = \frac{1}{2}$ the above expression is non-negative (at $p = 0$ from assumption).

We will show that $g(p)$ is concave in p when $p \in [0, \frac{1}{2}]$. Note that the function $g_1(x) = x(1 - x) \log \frac{1-x}{x}$ satisfies

$$g_1'(x) = (1 - 2x) \log \frac{1 - x}{x} - 1, \text{ and } g_1''(x) = -2 \log \frac{1 - x}{x} - \frac{1 - 2x}{x(1 - x)}.$$

Thus $g_1(x)$ is concave when $x \in [0, \frac{1}{2}]$, implying $g(p)$ is concave in p as desired.

$\square$

When $p = 0$, $H(p * f(u)) = H(f(u))$. Theorem 2 shows that the convexity of $H(p * f(u))$ directly follows its convexity at the endpoint $p = 0$. MGL follows from Theorem 2 obviously, because $H(H^{-1}(u)) = u$. Also, our argument simplifies the proof of MGL in [1].

Note that

$$\frac{df^{-1}}{du} = \frac{1}{f'(f^{-1}(u))}, \text{ and } \frac{d^2 f^{-1}}{du^2} = -\frac{f''(f^{-1}(u))}{[f'(f^{-1}(u))]^3}.$$

When $f(u)$ is replaced by $f^{-1}(u)$ in Theorem 2, $H(f^{-1}(u))$ is convex in u if and only if

$$(2) \qquad -f^{-1}(u)(1 - f^{-1}(u))\frac{f''(f^{-1}(u))}{f'(f^{-1}(u))} \log \frac{1 - f^{-1}(u)}{f^{-1}(u)} \geq 1.$$

Theorem 2 relies on the twice differentiability of $f(u)$. In the next theorem, we prove a strengthened version without this constraint.

Theorem 3. *For every $p \in [0,1]$, the function $H(p * f(u))$ is convex in u provided $H(f(u))$ is convex in u, where $f(u) : (a,b) \to [0, \frac{1}{2}]$.*

Though $f(u)$ is not twice differentiable, $f(u)$ is still convex by the convexity of $H(f(u))$. Since $f''(u)$ may not exist, we need an alternative method to deal with the convexity. Next, we state some instrumental results on convex function in Pollard [9] (Appendix C).

A convex function is always continuous and its one-sided derivatives always exist. For a convex function $f(x)$, denote its left-hand and right-hand derivatives by $f'_-(x)$ and $f'_+(x)$, respectively. Furthermore, both $f'_-(x)$ and $f'_+(x)$ are increasing; i.e.,

$$(3) \qquad f'_-(x_0) \le f'_-(x_1) \text{ and } f'_+(x_0) \le f'_+(x_1) \text{ for each } x_0 < x_1.$$

Conversely, when $f'_+(x)$ is increasing, $f(x)$ is convex.

Lemma 2. *If a real-valued function f has an increasing, real-valued right-hand derivative at each point of an open interval, then f is convex on that interval.*

Now, we prove Theorem 3.

Proof. As in Theorem 2, we can still assume $p \le \frac{1}{2}$. Hence $p * f(u) \le \frac{1}{2}$.

Since $f(u)$ is convex in u, for each $u_0 < u_1$,

$$f'_+(u_0) \le f'_+(u_1).$$

Let

$$s(u) := H(p * f(u)).$$

Then $s(u)$ is continuous in u. Since $H(x)$ is differentiable,

$$s'_+(u) = (1 - 2p)f'_+(u) \log \frac{1 - p * f(u)}{p * f(u)}.$$

To show $s'_+(u)$ is increasing in an interval, it is equivalent to show that $s'_+(u)$ is increasing locally; i.e.,

$$(4) \qquad\qquad s'_+(u_0) \le s'_+(u_1),$$

where $u_1 > u_0$ and $u_1 \to u_0$.

Since the right-hand derivative of $\log \frac{1 - p * f(u)}{p * f(u)}$ exists,

$$s'_+(u_1) = (1 - 2p)f'_+(u_1) \log \frac{1 - p * f(u_1)}{p * f(u_1)}$$

$$= (1 - 2p)f'_+(u_1)\left(\log\frac{1 - p * f(u_0)}{p * f(u_0)} - \frac{(1 - 2p)f'_+(u_0)(u_1 - u_0)}{(p * f(u_0))(1 - p * f(u_0))}\right).$$

To show (4), it is equivalent to show

$$(f'_+(u_1) - f'_+(u_0))\log\frac{1 - p * f(u_0)}{p * f(u_0)} - \frac{(1 - 2p)f'_+(u_0)f'_+(u_1)(u_1 - u_0)}{(p * f(u_0))(1 - p * f(u_0))} \geq 0.$$

That is

$$g_2(p) := (f'_+(u_1) - f'_+(u_0))(p * f(u_0))(1 - p * f(u_0))\log\frac{1 - p * f(u_0)}{p * f(u_0)}$$
$$- (1 - 2p)f'_+(u_0)f'_+(u_1)(u_1 - u_0) \geq 0.$$

Since $f'_+(u_1) \geq f'_+(u_0)$ and $p * f(u_0) \leq \frac{1}{2}$, $g_2(p)$ is also concave in p, similar to $g(p)$. Thus, the convexity of $H(p * f(u))$ follows from the convexity at the endpoints $p = 0$ and $p = \frac{1}{2}$, which completes the proof. □

It is easy to see that Theorem 2 and Theorem 3 still hold when $p \in [p_0, 1 - p_0]$, as long as $H(p_0 * f(u))$ is convex in u.

4. Application

As another example, we give a simple proof to the following result.

Theorem 4 (Claim 1 in [10]). *When $f = H(\frac{u}{2}) + H(\frac{1-u}{2})$, $H(p * f^{-1}(u))$ is convex in $u \in [f(0.06), f(0.5)]$ for every $p \in [0, \frac{1}{2}]$.*

Proof. Let $t = f^{-1}(u)$, $t \in [0.06, 0.5]$. Then

$$f'(u) = \frac{1}{2}\log\frac{1 - \frac{u}{2}}{\frac{u}{2}} - \frac{1}{2}\log\frac{1 - \frac{1-u}{2}}{\frac{1-u}{2}},$$
$$f''(u) = -\frac{1}{u(2 - u)} - \frac{1}{(1 - u)(1 + u)}.$$

By Theorem 2, it suffices to prove that $H(f^{-1}(u))$ is convex in u. By (2), we obtain that

$$\frac{\frac{1}{2}\log\frac{(2-t)(1-t)}{t(t+1)}}{\log\frac{1-t}{t}} \leq -\frac{2t^2 - 2t - 1}{(2 - t)(1 - t)}.$$

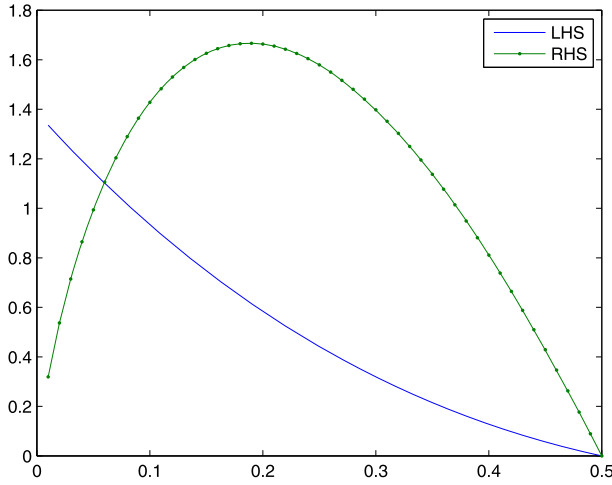Figure 1: Convexity of $f(u) = H(\frac{u}{2}) + H(\frac{1-u}{2})$.

By some algebra,

$$\frac{\log \frac{2-t}{t+1}}{\log \frac{1-t}{t}} \le \frac{7t - 5t^2}{(2-t)(1-t)}.$$

That is

$$(1-t)(2-t) \log \frac{2-t}{t+1} \le (7t - 5t^2) \log \frac{1-t}{t}.$$

Let

$$l(t) = (1-t)(2-t) \log \frac{2-t}{t+1}$$

and

$$r(t) = (7t - 5t^2) \log \frac{1-t}{t}.$$

The curves of the LHS ($l(t)$) and RHS ($r(t)$) are depicted in Fig. 1. By some algebra, we have

$$\frac{d^2 l(t)}{dt^2} = 2 \log \frac{2-t}{1+t} + \frac{3(3-2t)}{(1+t)(2-t)} + \frac{6}{(1+t)^2} \ge 0$$

and

$$\frac{d^2 r(t)}{dt^2} = -10 \log \frac{1-t}{t} - \frac{7-10t}{(1-t)t} - \frac{2}{(1-t)^2} \le 0.$$

When $t = 0.06$,

$$l(t) = 1.5902 \leq r(t) = 1.5958,$$

which completes the proof.   □

Acknowledgement

The author would like to thank Prof. Nair for providing this topic and thank Prof. Verdú and the anonymous reviewers for their very useful comments.

References

[1] A. D. Wyner and J. Ziv, "A theorem on the entropy of certain binary sequences and applications (Part I)," IEEE Trans. Inform. Theory, vol. 19, no. 6, pp. 769–772, Nov. 1973. MR0484773

[2] S. Shamai (Shitz) and A. D. Wyner, "A binary analog to the entropy power inequality," IEEE Trans. Inform. Theory, vol. 36, no. 6, pp. 1428–1430, Nov. 1990.

[3] H. Witsenhausen, "Entropy inequalities for discrete channels," IEEE Trans. Inform. Theory, vol. 20, no. 5, pp. 610–616, Sep. 1974. MR0360007

[4] R. Ahlswede and J. Korner, "On the Connection Between the Entropies of Input and Output Distributions of Discrete Memoryless Channels," in Proc. 5th Conf. Probability Theory, Braslov, 1974, pp. 13–23 (Academy Rep. Soc. Romania, Bucharest, 1977). MR0689182

[5] N. Chayat and S. Shamai (Shitz), "Extension of an entropy property for binary input memoryless symmetric channels," IEEE Trans. Inform. Theory, vol. 35, no. 5, pp. 1077–1079, Sep. 1989. MR1023246

[6] V. Jog and V. Anantharam, "The entropy power inequality and Mrs. Gerber's lemma for Abelian groups of order 2^n," Submitted to IEEE Trans. Inform. Theory. MR3225928

[7] A. D. Wyner, "A theorem on the entropy of certain binary sequences and applications (Part II)," IEEE Trans. Inform. Theory, vol. 19, no. 6, pp. 772–777, Nov. 1973. MR0484774

[8] C. Nair, "Capacity regions of two new classes of 2-receiver broadcast channels," IEEE Trans. Inform. Theory, vol. 56, no. 9, pp. 4207–4214, 2010. MR2807320

[9] D. Pollard, *A User's Guide to Measure Theoretic Probability*, Cambridge University Press, 2002. MR1873379

[10] C. Nair and Z. V. Wang, "On 3-receiver broadcast channels with 2-degraded message sets," IEEE International Symposium on Information Theory, pp. 1844–1848, 2009.

Fan Cheng
Institute of Network Coding
The Chinese University of Hong Kong, N.T.
Hong Kong
E-mail address: fcheng@inc.cuhk.edu.hk