The Method of Chaining

MA5249 Project Report Part 2

Pan Jing Bin
A0199300H

4 Introduction

In the first part of the report, we have started on our journey to develop a tool to control the supremum of random processes. While it is already able to establish non-trivial estimates in some situations, the tool is still somewhat unrefined and the bounds obtained are far from optimal. In the second part of the report, we will continue the journey to refine our tool.

We keep all notations in the first part of the report. To avoid confusion when referencing the results established in the first report, we will start the section numbering at 4.

5 The Chaining Method

In this section, we will fully develop the core of the tool: the chaining argument.

In the previous section, we used the Lipschitz property $|X_s - X_t| \lesssim d(s,t)$ to control the supremum of the remainder term in the sum

$$\sup_{t\in T} X_t \leq \sup_{t\in T} \Big\{X_t - X_{\pi(t)}\Big\} + \sup_{t\in T} X_{\pi(t)}.$$

In general, such a condition is very restrictive as the Lipschitz bound C needs to be sufficiently large to account for the worst case. In many situations, the typical size of the increments $X_s - X_t$ is much smaller than in the worst case scenario.

However, if we only require the Lipschitz property to hold in probability instead of almost surely, then another problem arises: we cannot directly control the remainder term. This is because even if each variable is "typically" small, since we have to control the supremum of many such variables, their total magnitude can be significantly larger. For example, the maximum of n i.i.d standard normal variables is asymptotically $\gtrsim \log n$. Hence the problem of controlling the remainder term is essentially of the same type as the original problem.

Nonetheless, we still expect the task of estimating the remainder to be easier than the original task as the size of each variable in the remainder term tends to be smaller. The key idea here is to shrink the remainder term even further by repeating the original argument with a finer net. For example, if N' is an $\epsilon/2$-net, then we can estimate

$$\sup_{t\in T} \big\{X_t - X_{\pi(t)}\big\} \leq \sup_{t\in T} \big\{X_t - X_{\pi'(t)}\big\} + \sup_{t\in T} \big\{X_{\pi'(t)} - X_{\pi(t)}\big\}.$$

This process can be repeated any number of times to yield

$$\sup_{t\in T} X_t \le \sup_{t\in T} \overbrace{\{X_t - X_{\pi_n(t)}\}}^{\sim 2^{-n}} + \sum_{k=1}^{n} \sup_{t\in T} \overbrace{\{X_{\pi_k(t)} - X_{\pi_{k-1}(t)}\}}^{\sim 2^{-k}} + \sup_{t\in T} X_{\pi_0(t)}. \tag{1}$$

For this approximation to work properly, we need to control the telescoping series and ensure the the remainder term $X_t - X_{\pi_n(t)}$ vanish as $n \to \infty$. This motivates us to consider these two definitions.

Definition 5.1. (Subgaussian process) A random process $\{X_t\}_{t\in T}$ on the metric space (T, d) is **subgaussian** if $\mathbb{E}[X_t] = 0$ and

$$\mathbb{E}\big[e^{\lambda(X_s - X_t)}\big] \le e^{\frac{\lambda^2 d(s,t)^2}{2}} \qquad \text{for all } s, t \in T \text{ and } \lambda \ge 0.$$

Remark 5.2. By comparing Taylor series and letting $\lambda \to 0$, the above definition already implies that $\mathbb{E}[X_s - X_t] = 0$ for all $s, t \in T$. Thus the assumption $\mathbb{E}[X_t] = 0$ is simply a convenient normalization. In the next section, we will generalise the techniques developed to processes with non-trivial mean behavior $t \mapsto \mathbb{E}[X_t]$.

Definition 5.3. (Separable process) A random process $\{X_t\}_{t\in T}$ is **separable** if there exists a countable subset $T_0 \subseteq T$ and an event E of probability 1 such that for all $\omega \in E$ and $t \in T$, there exists a sequence $(t_k)_{k=1}^{\infty}$ in T_0 satisfying

$$\lim_{k\to\infty} X_{t_k}(\omega) = X_t(\omega).$$

Remark 5.4. The separability assumption implies that

$$\sup_{x\in T} X_t = \sup_{x\in T_0} X_t \qquad \text{a.s.}$$

Thus all issues pertaining to measurability can be addressed by passing to a countably dense subset.

With the subgaussian and separability assumptions, we now have all the ingredients necessary to implement the chaining argument.

Theorem 5.5. (Dudley) Let $\{X_t\}_{t\in T}$ be a separable subgaussian process on the metric space (T, d). Then we have the following estimate:

$$\mathbb{E}\left[\sup_{t\in T} X_t\right] \le 6 \sum_{k\in\mathbb{Z}} 2^{-k} \sqrt{\log N(T, d, 2^{-k})}.$$

Proof. As mentioned in our previous discussion, it suffices to prove the statement when $|T|$ is finite. The separability of the process can then be used to lift the restriction. Let k_0 be the largest integer such that $2^{-k_0} \ge \operatorname{diam}(T)$. Then any singleton $N_{k_0} = \{t_0\}$ is trivially a 2^{-k_0}-net and we start our chaining process from there. For each $k > k_0$, let N_k be a 2^{-k}-net satisfying $|N_k| = N(T, d, 2^{-k})$. Running the chaining argument up to the scale 2^{-n} yields

$$\mathbb{E}\left[\sup_{t\in T} X_t\right] \le \mathbb{E}\left[\sup_{t\in T}\left\{X_t - X_{\pi_n(t)}\right\}\right] + \sum_{k=k_0+1}^{n} \mathbb{E}\left[\sup_{t\in T}\left\{X_{\pi_k(t)} - X_{\pi_{k-1}(t)}\right\}\right] + \mathbb{E}[X_{t_0}].$$

By assumption, $\mathbb{E}[X_{t_0}] = 0$ and so the first term vanishes. By choosing n sufficiently large, we have $N_n = T$ and so the last term vanishes as well. Each of the terms in the summation is a supremum over a finite set consisting of at most $|N_k||N_{k-1}| \le |N_k|^2$ elements. Furthermore,

$$d\big(\pi_k(t), \pi_{k-1}(t)\big) \le d\big(\pi_k(t), t\big) + d\big(t, \pi_{k-1}(t)\big) \le 3 \times 2^{-k}.$$

Thus each $X_{\pi_k(t)} - X_{\pi_{k-1}(t)}$ is $(3 \times 2^{-k})^2$-subgaussian. By Lemma 2.1, we have the bound

$$\mathbb{E}\left[\sup_{t\in T}\left\{X_{\pi_k(t)} - X_{\pi_{k-1}(t)}\right\}\right] \leq 6 \times 2^{-k}\sqrt{\log N(T,d,2^{-k})}.$$

Taking summation, we obtain

$$\mathbb{E}\left[\sup_{t\in T} X_t\right] \leq 6\sum_{k=k_0+1}^{\infty} 2^{-k}\sqrt{\log N(T,d,2^{-k})} = 6\sum_{k\in\mathbb{Z}} 2^{-k}\sqrt{\log N(T,d,2^{-k})}.$$

□

Dudley's theorem can also be reformulated in terms of an integral. In analogy with information theory, the quantity $\log N(T,d,\epsilon)$ is called the **metric entropy** and the integral is called the **entropy integral**.

Corollary 5.5.1. (Entropy Integral) Let $\{X_t\}_{t\in T}$ be a separable subgaussian process on the metric space (T,d). Then we have the following estimate:

$$\mathbb{E}\left[\sup_{t\in T} X_t\right] \leq 12\int_0^{\infty} \sqrt{\log N(T,d,\epsilon)}\ d\epsilon.$$

Proof. A direct computation gives

$$\begin{aligned}\sum_{k\in\mathbb{Z}} 2^{-k}\sqrt{\log N(T,d,2^{-k})} &= 2\sum_{k\in\mathbb{Z}}\int_{2^{-k-1}}^{2^{-k}} \sqrt{\log N(T,d,2^{-k})}\ d\epsilon \\ &\leq 2\sum_{k\in\mathbb{Z}}\int_{2^{-k-1}}^{2^{-k}} \sqrt{\log N(T,d,\epsilon)}\ d\epsilon = 2\int_0^{\infty} \sqrt{\log N(T,d,\epsilon)}\ d\epsilon.\end{aligned}$$

where the second last inequality follows from the observation that the map $\epsilon \mapsto N(T,d,\epsilon)$ is non-increasing. □

To demonstrate the significance of this improvement, let us revisit Example 3.6.

Example 5.6. (Wasserstein law of large numbers revisited) By another routine application of the Azuma-Hoeffding inequality, we see that

$$\mathbb{E}\left[e^{\lambda(X_f - X_g)}\right] \leq e^{\frac{\lambda^2\|f-g\|_\infty^2}{2n}}$$

for any $f, g \in \mathrm{Lip}([0,1])$ and so the process $\{X_f\}_{f\in\mathcal{F}}$ is subgaussian with respect to the metric $d(f,g) = n^{-1/2}\|f-g\|_\infty$. Using Corollary 5.5.1 and Lemma 3.8, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{f\in\mathcal{F}} X_f\right] &\leq 12\int_0^{\infty} \sqrt{\log N\left(\mathcal{F}, n^{-1/2}\|\cdot\|_\infty, \epsilon\right)}\ d\epsilon = 12\int_0^{\infty} \sqrt{\log N\left(\mathcal{F}, \|\cdot\|_\infty, n^{1/2}\epsilon\right)}\ d\epsilon \\ &= \frac{12}{\sqrt{n}}\int_0^{\infty} \sqrt{\log N\left(\mathcal{F}, \|\cdot\|_\infty, \epsilon\right)}\ d\epsilon \leq \frac{12}{\sqrt{n}}\int_0^{1/2} \sqrt{\frac{c}{\epsilon}}\ \ d\epsilon.\end{aligned}$$

Since the integral converges, we obtain the improved estimate

$$\mathbb{E}\left[\sup_{f\in\mathcal{F}} X_f\right] \lesssim n^{-1/2}.$$

Using the central limit theorem, it is easy to see that the upper bound of $n^{-1/2}$ is asymptotically optimal.

The general principle described in (1) allows us to establish upper bounds on the tail probability in an analogous manner by using Lemma 2.2 in place of Lemma 2.1.

Theorem 5.7. (Chaining tail inequality) Let $\{X_t\}_{t\in T}$ be a separable subgaussian process on the metric space (T,d). Then for all $t_0 \in T$ and $x \geq 0$, we have the estimate

$$P\left(\sup_{t\in T}\{X_t - X_{t_0}\} \geq C\int_0^\infty \sqrt{\log N(T,d,\epsilon)}\ d\epsilon + x\right) \leq Ce^{-\frac{x^2}{C\cdot \mathrm{diam}(T)^2}}$$

for some universal constant $C \in \mathbb{R}_{>0}$.

Corollary 5.7.1. (Local chaining inequality) Let $\{X_t\}_{t\in T}$ be a separable subgaussian process on the metric space (T,d). Then for all $x, \delta \in \mathbb{R}_{\geq 0}$, we have the estimate

$$P\left(\sup_{\substack{s,t\in T\\ d(s,t)\leq\delta}} \{X_s - X_t\} \geq C\int_0^\delta \sqrt{\log N(T,d,\epsilon)}\ d\epsilon + x\right) \leq Ce^{-\frac{x^2}{C\delta^2}}$$

for some universal constant $C \in \mathbb{R}_{>0}$.

Proof. (Sketch) For each $s,t \in T$, define $\widetilde{X}_{s,t} = X_s - X_t$. Define $\widetilde{T} := \{(s,t) \in T\times T : d(s,t) \leq \delta\}$. It is not difficult to show that the metric

$$\widetilde{d}\big((s,t),(u,v)\big) = \min\left\{2^{1/2}\sqrt{d(s,u)^2 + d(t,v)^2},\ 2\delta\right\}$$

turns $\{\widetilde{X}_{s,t}\}_{(s,t)\in\widetilde{T}}$ into a subgaussian process. The conclusion follows by applying the Theorem 5.7 together with the observations that $\mathrm{diam}(\widetilde{T}) \leq 2\delta$ and $N(\widetilde{T},\widetilde{d},2\epsilon) \leq N(T,d,\epsilon)^2$. □

6 Penalization and the slicing method

In the previous section, we have completed the core of our tool, which is the chaining argument. Nevertheless, there is still much room for improvement. A key assumption that we have made is that the random process $\{X_t\}_{t\in T}$ is centered at $\mathbb{E}[X_t] = 0$. To generalise the technique to processes with non-trivial mean behavior $t \mapsto \mathbb{E}[X_t]$, we decompose our process into

$$\sup_{t\in T} X_t = \sup_{t\in T}\{\mathbb{E}[X_t] + Z_t\}$$

where the fluctuations $\{Z_t\}_{t\in T}$ are assumed to form a subgaussian process. This is known as an additive **penalized supremum**. More generally, we may also consider multiplicative penalized supremums such as

$$\sup_{s,t\in T}\frac{X_s - X_t}{\rho(s,t)}.$$

The method of slicing is a way to generalise the method of chaining to these cases. We first choose a sequence $(a_k)_{k=1}^\infty$ decreasing to 0 and decompose the supremum into "slices"

$$\begin{aligned}
P\left(\sup_{s,t\in T}\left\{\frac{X_s - X_t}{\rho(s,t)}\right\} \geq x\right) &= P\left(\sup_{k\geq 1}\left\{\sup_{\alpha_k\leq\rho(s,t)\leq\alpha_{k-1}}\left\{\frac{X_s - X_t}{\rho(s,t)}\right\}\right\} \geq x\right)\\
&\leq \sum_{k=1}^\infty P\left(\sup_{\alpha_k\leq\rho(s,t)\leq\alpha_{k-1}}\left\{\frac{X_s - X_t}{\rho(s,t)}\right\} \geq x\right)\\
&\leq \sum_{k=1}^\infty P\left(\sup_{\rho(s,t)\leq\alpha_{k-1}}\{X_s - X_t\} \geq \alpha_k x\right).
\end{aligned}$$

Each term in the summation is now the tail of the supremum of a subgaussian process without penalty. However, the penalty still appears implicitly, as it determines the subset of the index set over which the supremum is taken in each term of the sum. If α_k is small, then the supremum is taken over a smaller set so we expect the probability to decrease. However, the threshold $\alpha_k x$ also decreases and so the probability should also (somewhat) increase. Hence some careful thought must be given in order to choose the sequence $(\alpha_k)_{k=1}^{\infty}$ to ensure that we obtain a satisfactory convergent sum.

In general, finding such a sequence $(\alpha_k)_{k=1}^{\infty}$ requires some finesse and depends on the specific context of the problem. In this report, we will illustrate one such example.

Example 6.1. (Modulus of continuity) Another quantity of subgaussian processes that is of significant interest in its own right is the "degree of smoothness" of the map $t \mapsto X_t$.

Definition 6.2. An increasing function $\omega : \mathbb{R} \to \mathbb{R}$ is a **modulus of continuity** for a random process $\{X_t\}_{t\in T}$ on the metric space (T, d) if $\omega(0) = 0$ and there exists a random variable K such that

$$X_s - X_t \leq K\omega(d(s,t)) \qquad \text{for all } s, t \in T.$$

Here, we will derive an explicit modulus of continuity for subgaussian processes.

Theorem 6.3. Let $\{X_t\}_{t\in T}$ be a separable subgaussian process on the metric space (T, d). Assume that there exist constants $c, q \in \mathbb{R}_{>0}$ such that $N(T, d, \epsilon) \geq (c/\epsilon)^q$ for all $\epsilon \in \mathbb{R}_{>0}$. Then

$$\omega(\delta) := \int_0^{\delta} \sqrt{\log N(T,d,\epsilon)}\, d\epsilon$$

is a modulus of continuity for $\{X_t\}_{t\in T}$.

Proof. Showing that ω is a modulus of continuity is equivalent to proving that

$$\sup_{s,t\in T} \frac{X_s - X_t}{\omega(d(s,t))} < \infty \quad \text{a.s.}$$

We have hence reformulated our problem as the supremum of a random process. Let $D = \operatorname{diam}(T)$. Applying the slicing argument with $\alpha_k := \omega(2^{-k}D)$, we get

$$\begin{aligned} P\left(\sup_{s,t\in T}\left\{\frac{X_s - X_t}{\omega(d(s,t))}\right\} \geq 2(C+x)\right) &\leq \sum_{k=1}^{\infty} P\left(\sup_{d(s,t)\leq 2^{-k+1}D} \{X_s - X_t\} \geq 2\omega(2^{-k}D)(C+x)\right) \\ &\leq \sum_{k=1}^{\infty} P\left(\sup_{d(s,t)\leq 2^{-k+1}D} \{X_s - X_t\} \geq \omega(2^{-k+1}D)(C+x)\right) \end{aligned}$$

where the last line follows from the observation that $\omega(2^{-k+1}D) \leq 2\omega(2^{-k}D)$. By Corollary 5.7.1,

$$\begin{aligned} &P\left(\sup_{s,t\in T}\left\{\frac{X_s - X_t}{\omega(d(s,t))}\right\} \geq 2(C+x)\right) \\ &\leq \sum_{k=1}^{\infty} P\left(\sup_{d(s,t)\leq 2^{-k+1}D} \{X_s - X_t\} \geq (C+x)\int_0^{2^{-k+1}D} \sqrt{\log N(T,d,\epsilon)}\, d\epsilon\right) \\ &\leq \sum_{k=1}^{\infty} C\exp\left(-\frac{x^2}{C}\left(\frac{1}{2^{-k+1}D}\int_0^{2^{-k+1}D} \sqrt{\log N(T,d,\epsilon)}\, d\epsilon\right)^2\right) \\ &\leq \sum_{k=1}^{\infty} C\exp\left(-\frac{x^2}{C}\log N\big(T,d,2^{-k+1}D\big)\right), \end{aligned}$$

where we have again used the fact that $\epsilon \mapsto \sqrt{\log N(T,d,\epsilon)}$ is non-increasing for the last inequality. Now we use the technical assumption $N(T,d,\epsilon) \geq (c/\epsilon)^q$ to upper bound the sum by a convergent geometric series. After simplifying, we obtain an expression of the form

$$P\left(\sup_{s,t\in T}\left\{\frac{X_s - X_t}{\omega(d(s,t))}\right\} \geq 2(C+x)\right) \lesssim Ae^{-\frac{x^2}{A}}.$$

This completes the proof. □

7 Lower Bounds for Gaussian Processes

At this point, there is much reason to be skeptical about the general effectiveness of the method of chaining: at first sight the method appears quite crude, being at its core little more than just a conveniently organized union bound. To understand when the chaining method give sharp bounds, we must supplement our upper bounds with corresponding lower bounds.

Establishing lower bounds is generally considered to be a much harder task than establishing upper bounds and tackling such a problem at the level of generality of subgaussian processes is almost impossible. Some additional structure, such as independence, is usually required to obtain meaningful estimates. In this final section, we will restrict our attention to only Gaussian processes, one of the most important and well-studied processes in probability theory.

Definition 7.1. (Gaussian process) The random process $\{X_t\}_{t\in T}$ is called a **(centered) Gaussian process** if for all $n \in \mathbb{Z}_{\geq 1}$ and indices $t_1, \cdots, t_n$, the random variables $\{X_{t_1}, \cdots, X_{t_n}\}$ are centered (i.e. $\mathbb{E}[X_{t_j}] = 0$ for each j) and jointly Gaussian.

Remark 7.2. For a Gaussian process $\{X_t\}_{t\in T}$, we have

$$\mathbb{E}\big[e^{\lambda(X_t - X_s)}\big] = \exp\left(\frac{\lambda^2 \mathbb{E}\big[|X_t - X_s|^2\big]}{2}\right).$$

Thus a Gaussian process induces a canonical metric on the index set T.

Definition 7.3. (Natural distance) A Gaussian process $\{X_t\}_{t\in T}$ is subgaussian on (T,d) under the natural distance $d(s,t) := \mathbb{E}\big[|X_s - X_t|^2\big]^{1/2}$.

To establish lower bounds, we once again revisit the finite case. With the additional i.i.d Gaussian assumption, the bound given in Lemma 2.1 is in fact sharp.

Lemma 7.4. If $X_1, \cdots, X_n$ are i.i.d $\mathcal{N}(0,\sigma^2)$ random variables, then

$$c\sqrt{\sigma^2 \log n} \leq \mathbb{E}\left[\max_{i\leq n} X_i\right] \leq \sqrt{2\sigma^2 \log n}$$

for some universal constant $c \in \mathbb{R}_{\geq 0}$.

Throughout the development of our theory, a key principle that we have repeatedly used is that two random variables X_s and X_t are strongly dependent when s and t are close together. Conversely, we expect X_s and X_t to be nearly independent when s and t are far apart since we do not have any control over their relationship. However, we have never used any form of independence in our proofs so far. The theory still holds even if the far away points are strongly dependent. On the other hand, one can only expect these bounds to be sharp if our intuition do hold. To leverage on the relationship of well-separated points, we will take advantage of the following comparison inequality.

Theorem 7.5. (Slepian-Fernique) Let $X \sim \mathcal{N}(0, \Sigma_X)$ and $Y \sim \mathcal{N}(0, \Sigma_Y)$ be n-dimensional Gaussian vectors. Suppose that we have

$$\mathbb{E}\big[|X_i - X_j|^2\big] \geq \mathbb{E}\big[|Y_i - Y_j|^2\big] \quad \text{for all } i, j \in \{1, \cdots, n\}.$$

Then

$$\mathbb{E}\left[\max_{1\leq k\leq n} X_k\right] \geq \mathbb{E}\left[\max_{1\leq k\leq n} Y_k\right].$$

This theorem allows us to lower bound the supremum of (possibly dependent) Gaussian variables that are far apart by the supremum of independent Gaussian variables. Assuming this classical result, the next theorem follows as an easy corollary.

Theorem 7.6. (Sudakov) For a Gaussian process $\{X_t\}_{t\in T}$, we have the lower bound

$$\mathbb{E}\left[\sup_{t\in T} X_t\right] \geq \widetilde{c} \sup_{\epsilon>0} \epsilon\sqrt{\log N(T, d, \epsilon)}$$

for a universal constant $\widetilde{c} \in \mathbb{R}_{\geq 0}$.

Proof. Fix $\epsilon > 0$ and an ϵ-packing $N = \{t_1, \cdots, t_n\}$ of T, where $n = D(T, d, \epsilon)$.

Let $X = \{X_{t_k}\}_{k=1}^n$. Further let $Y = \{Y_k\}_{k=1}^n$ be i.i.d $\mathcal{N}(0, \epsilon^2/2)$ random variables. Then

$$\mathbb{E}\big[|X_s - X_t|^2\big] = d(s,t)^2 \geq \epsilon^2 = \mathbb{E}\big[|Y_s - Y_t|^2\big] \qquad \text{for all } s, t \in N, s \neq t.$$

Hence we may use Theorem 7.5 to lower bound the supremum of $\{X_t\}_{t\in T}$ by the independent system $\{Y_k\}_{k=1}^n$. Together with Lemmas 3.5 and 7.4, we obtain

$$\mathbb{E}\left[\sup_{t\in T} X_t\right] \geq \mathbb{E}\left[\max_{1\leq k\leq n} X_{t_k}\right] \geq \mathbb{E}\left[\max_{1\leq k\leq n} Y_k\right] \geq \widetilde{c}\epsilon\sqrt{\log D(T, d, \epsilon)} \geq \widetilde{c}\epsilon\sqrt{\log N(T, d, \epsilon)}$$

for some universal constant $\widetilde{c} \in \mathbb{R}_{>0}$ (possibly different from the universal constant in Lemma 7.4). □

Hence we obtain both upper and lower bounds on the supremum of $\{X_t\}_{t\in T}$ that are defined in terms of the same quantity

$$\sup_{\epsilon>0} \epsilon\sqrt{\log N(T, d, \epsilon)} \lesssim \mathbb{E}\left[\sup_{t\in T} X_t\right] \lesssim \int_0^\infty \sqrt{\log N(T, d, \epsilon)}\, d\epsilon.$$

8 Conclusion

While the theory that we have developed so far is already powerful enough to establish non-trivial asymptotic bounds in many situations, we have only just begun to scratch the surface. In particular, the usage of many crude union bounds in the process means that resulting estimates are usually not sharp. The question of how to improve upon the tool even further continues to generate much research to this day. One of the most successful instances is Talagrand's work in generalising the chaining argument to arbitrary measures (known as majorizing measures). The end result is now known as generic chaining. The tool is so remarkably effective that it is able to produce estimates that are essentially optimal in many situations.

9 References

M. Talagrand. *The generic chaining. Upper and lower bounds of stochastic processes*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005.

R. van Handel. *Probability in High Dimension*. Course notes, Princeton University, 21 Dec 2016. Retrieved from https://web.math.princeton.edu/∼rvan/APC550.pdf