

3 Ways to Stop Wasting Your Claude Limit

Three free habits that stop the burn

Copy the wording exactly. If you only do one, do fix 3. It pays you back on every chat after it.

The limits feel tighter. A lot of that burn is self-inflicted.

1. PASTE TEXT, NOT JUNK

Here is the full text of [document name]. Read it, then [what you want]. If anything looks garbled from the copy and paste, tell me which part instead of guessing at it.

Use it for: Reports, contracts, transcripts, decks.

How: open the PDF, select all, copy, and paste the plain text into the chat instead of attaching the file. **Why it works:** Anthropic's documentation says a PDF gets its text extracted and every page converted into an image, and that each page typically runs 1,500 to 3,000 tokens before that image cost lands on top. The formatting, the logos, the scanned letterhead: none of that was the part you needed. If you have to keep the file, split it and send only the pages that matter.

2. HANDOFF DOC, THEN FRESH CHAT

Stop where we are. Write me a handoff document I can paste into a brand new chat. Include, in this order: 1) the goal in one sentence, 2) the constraints and requirements we have settled on, 3) what we already tried and why it did not work, 4) the current state of the work, 5) the exact next step. Plain text, no preamble, no summary of this conversation. If something is still unresolved, write "unresolved" instead of guessing.

Use it for: Any chat where you are on the fourth round of corrections.

Then: open a new chat, paste the handoff, and start with "Picking this up from a previous session. Here is the handoff. Start at the next step." **Why it works:** Anthropic's guidance is direct about it. Longer conversations that trigger automatic context management consume more of your usage limit, and the recommendation is to start a new conversation if you are approaching your limit in a long chat. Take the good part with you and leave the arguing behind.

3. SAY IT ONCE

Standing instructions for this project.

Who I am: [role, business, who I serve]

What we work on here: [the recurring job]

Output format: [bullets, tables, word count, no preamble]

Tone: [plain, direct, no hype]

Always: [the two or three rules you repeat most]

Never: [the two or three things you keep correcting]

When you do not have enough information, ask me one question instead of guessing.

Use it for: Anything you explain more than twice.

How: paste it into the project instructions field of a Claude Project and fill in the brackets. Every new chat inside that project starts already knowing. Projects are available on free accounts too, capped at five. Files you upload to a project are cached, so only the new parts count against your limit each time you reference them. **On Claude Code:** the same idea lives in a CLAUDE.md file at your project root, loaded at the start of every session. Run /init to generate a starter, keep it under 200 lines, and write rules specific enough to check ("use 2-space indentation" beats "format code properly").

4 bonus habits, from Anthropic's own usage guidance

- **Batch your questions.** Related tasks go in one message, not five.
- **Point back, do not re-paste.** "As mentioned earlier" beats pasting the brief in again.
- **Upload once, per project.** Project files get cached. Re-attaching the same doc to every chat is paying twice.
- **Watch the bar.** Settings, then Usage. Progress bars for your five-hour session and your week.

Want the rest?

Every free guide from the videos lives on one page: keybuilds.ai/start

The paid packs go deeper, full prompt libraries and workflow chains: go.keybuilds.ai